Mathematisches Forschungsinstitut Oberwolfach

Report No. 20/2015

DOI: 10.4171/OWR/2015/20

Copulae: On the Crossroads of Mathematics and Economics

Organised by
Xiaohong Chen, New Haven
Wolfgang Karl Härdle, Berlin
Piotr Jaworski, Warsaw
Johanna G. Nešlehová, Montréal

12 April – 18 April 2015

Abstract. The central focus of the workshop was on copula theory as well as applications to multivariate stochastic modelling. The programme was intrinsically interdisciplinary and represented areas with much recent progress. The workshop included talks and dynamic discussions on construction, estimation and various applications of copulas to finance, insurance, hydrology, medicine, risk management and related fields.

Mathematics Subject Classification (2010): 62H05, 60E05, 91B30, 91G40, 91G70, 62G05.

Introduction by the Organisers

The workshop "Copulae: On the Crossroads of Mathematics and Economics", organized by Xiaohong Chen (Yale University, New Haven, CT), Wolfgang Karl Härdle (Humboldt-Universität zu Berlin), Piotr Jaworski (University of Warsaw) and Johanna G. Nešlehová (McGill University, Montréal, QC), was held 12–18 April, 2015. The meeting was very well attended with 26 participants and broad representation from Europe, North America and Asia, as well as varied background including statistics, probability, econometrics, financial mathematics and uncertainty modelling. In total, 23 research talks of varying length were delivered by both leading experts and young researchers in the field. In addition, an introductory lecture on copula-based dependence modelling was given by Christian Genest at a joint session with the parallel workshop on "Mathematical Theory of Water Waves". A summary of this talk will be available within the Oberwolfach snapshot series (http://www.mfo.de/math-in-public/snapshots).

The central focus of the workshop was copula theory and its applications to multivariate stochastic modelling. Copulas are mathematical objects that capture the dependence structure among random variables and offer great flexibility in building multivariate stochastic models. Since their discovery in the early 1950's, copulas have led to a much better understanding of stochastic dependence; they have also allowed modellers to break away from the multivariate Gaussian distribution, which generally underestimates the probability of joint extreme risks. Copula-based dependence models are rapidly gaining popularity and are becoming indispensable tools, e.g., in biostatistics, econometrics, hydrology, finance, insurance, and risk management. For example, they are well suited for the modelling of market, credit and operational risk, for risk aggregation and portfolio selection.

The aim of the workshop was to bring together researchers in mathematical statistics, probability, econometrics, actuarial science, and risk management to discuss recent developments and to address challenges in the field. The programme was intrinsically interdisciplinary and represented areas with much recent progress; the themes of the talks centred around the construction, estimation and applications of copulas to multivariate modelling in finance, insurance and other fields. The meeting generated an intensive exchange of ideas and lively discussions. Smaller groups of participants often formed during the breaks and in the evenings for additional deliberations. A number of important research contacts were made which will no doubt spark new collaborative research projects in the near future.

The excellent scientific programme was spiced up by two scheduled social activities. On Wednesday afternoon, good use was made of the splendid weather for a hike to St. Roman and the traditional Black Forest Cake. The second activity was the Friday afternoon visit to the Museum for Minerals and Mathematics in Oberwolfach Kirche. Many participants were on their first visit to Oberwolfach and left very impressed with the experience. There was a strong consensus that the workshop on "Copulae: On the Crossroads of Mathematics and Economics" should become a regular Oberwolfach event.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-1049268, "US Junior Oberwolfach Fellows".

Workshop: Copulae: On the Crossroads of Mathematics and Economics

Table of Contents

Abstracts

Copulas in financial risk management

Alexander J. McNeil

In my contribution to the workshop I give an overview of the main applications of copulas in the area of quantitative financial risk management. I begin with a summary of some basic concepts in copula modelling, including shared frailty models, Archimedean copulas and factor copulas, before going on to identify three main applications: copulas in portfolio credit risk models; copulas in risk aggregation; copulas in econometric models of financial returns for market risk modelling. I argue that the ideas for these applications came in many instances from actuarial science where there is a longer tradition of using copulas to model dependent lifetimes and dependent multi-line insurance losses.

I devote the majority of the talk to credit risk where I distinguish between models for calculating measures of portfolio credit risk like VaR (Value-at-Risk) under the real-world probability measure P and models for valuing portfolio credit risk products (such as CDOs) under a risk-neutral measure Q. I show how the former models descend from a multivariate version of Merton's model, which implies a Gaussian copula for the latent variables that drive credit risk. I then discuss how standard industry calibration approaches are subject to considerable model risk due to uncertainty about the "true underlying copula". This observation was made by Frey, McNeil & Nyeler [1] some years before the 2007-09 financial crisis.

I go on to discuss the approach proposed by Li [2] to valuing CDOs that became extremely popular before the crisis. Inspired by actuarial models of correlated lifes and the so-called broken-heart syndrome, Li suggested the use of the Gaussian copula to model correlated times to default under Q. The disastrous consequences of the widespread adoption of this approach and the neglect of the model risk involved have been well documented.

The aggregation problem is a classic example of the use of the converse of Sklar's theorem to build joint models of risk across an enterprise from marginal models and copulas; the resulting models are used for capital calculation and allocation. The great difficulty lies in selecting and calibrating an appropriate copula, particularly when data are scarce. I highlight a method by Arbenz, Hummel and Mainik [4] which attempts to aggregate capital over a tree structure when marginal information is fully specified but the dependence is only specified through copulas at branching nodes of the tree.

Finally I show how popular econometric models for financial returns such as the DCC model of Engle [3] are actually models with a time-varying conditional Gaussian copula, Student t or other normal variance-mixture copula. The parameters that vary with time are the correlation parameters of the copula. This precludes, for example, a time-varying degree-of-freedom parameter in the t copula. I survey literature that tries to extend the class of econometric models to

allow dynamic behaviour of further parameters of the copula; see, for example, Fan and Patton [5].

References

[1] R. Frey, A. J. McNeil, and M. Nyfeler, *Copulas and credit models*, Risk, **14(10)** (2001), 111–114.

[2] D. Li, *On default correlation: A copula function approach*, Journal of Fixed Income, **9** (2000), 43–54.

[3] R. F. Engle, *Dynamic conditional correlation — a simple class of multivariate GARCH models*, Journal of Business & Economic Statistics, **20** (2002), 339–350.

[4] P. Arbenz, C. Hummel, and G. Mainik, *Copula based hierarchical risk aggregation though sample reordering*, Insurance: Mathematics and Economics, **51** (2012), 122–133.

[5] Y. Fan, and A. J. Patton, *Copulas in econometrics*, Annual Review of Economics, **6** (2014), 179–200.

Vine copulas: Introduction and recent advances

Claudia Czado

(joint work with K. Aas, E.C. Brechmann, L. Gruber, K. Hendrich, O. Kähm, M. Killiches and T. Nagler)

General dependency modeling is now a days often facilitated using a copula approach. Therefore different marginal distributions can be combined with a copula to form a joint distribution. The classes of multivariate copulas in more than two dimensions was limited to the class of elliptical and Archimedean copulas before the arrival of vine copulas. These pair copula constructions (PCC) combine bivariate copulas together with conditioning arguments to give a multivariate copula ([2]). A PCC consists out of a vine tree structure identifying the conditioning variables, the pair copula families for each edge in the vine tree structure and the associated pair copulas. These building blocks can be chosen arbitrary and separately, thus allowing for enormous modeling flexibility. This was realized in [1]. A first introduction can be found [4] including stepwise parameter estimation procedures. Choosing the vine tree structure is a challenging problem and current approaches are sumarized in [5]. The PCC approach is illustrated in a recent application to systemic risk stress testing where the stressing involved a single company ([3]) or multiple companies ([6]).

Recent advances include non simplified vines, truncated vines, vines with cross serial dependence, time varying/regime switching vines, discrete and discrete/continuous vines, non Gaussian DAG's using pair copula constructions, vines with non parametric copulas, acceleration of MCMC algorithms using vines, factor copula models, high dimensional goodness of fit for vines and spatial and spatial/temporal vines. The corresponding references can be found on the web page `vine-copula.org` as well as a multitude of applications in financial risk management, hydrology, machine learning, health, environmental science, chemistry,

energy and marketing. As examples of these advances the finite sample block maxima of vines ([7]) and multivariate kernel density estimation based on simplified vines ([8]) were illustrated.

References

[1] K. Aas, C. Czado, A. Frigessi, and H. Bakken, *Pair-copula constructions of multiple dependence*, Insurance, Mathematics and Economics **44** (2009), 182–198.

[2] T. Bedford, R.M. Cooke,*Pair-copula constructions of multivariate copulas Vines - a new graphical model for dependent random variables*, Annals of Statistics **30(4)** (2002), 1031–1068.

[3] E.C. Brechmann, K. Hendrich, and C. Czado , *Conditional copula simulation for systemic risk stress testing*, Insurance: Mathematics and Economics **53(3)** (2013), 829-839.

[4] C. Czado,*Pair-copula constructions of multivariate copulas*, Copula theory and its applications. Springer Berlin Heidelberg (2010), 93–109.

[5] C. Czado, E.C. Brechmann and L. Gruber, *Selection of vine copulas*, Copulae in Mathematical and Quantitative Finance. Springer Berlin Heidelberg, (2013), 17–37.

[6] O. Kähm, *Assessing system relevance of financial institutions using pair-copula constructions for modeling*, Master's thesis, Zentrum Mathematik, Technische Universität München (2014).

[7] M. Killiches, C. Czado, *Block-maxima of vines*, In D. Dey and J. Yan (Eds.), Extreme Value Modeling and Risk Analysis: Methods and Applications. Chapman & HallCRC Press, (2015).

[8] T. Nagler and C. Czado (2015). *Evading the curse of dimensionality in multivariate kernel density estimation with simplifed vines*, `arXiv:1503.03305v1 [stat.ME]`, 2015.

Copula calibration

Johanna Ziegel

(joint work with Tilmann Gneiting)

Probabilistic predictions of a future event $\mathbf{Y} \in \mathbb{R}^d$ are predictions that specify the entire distribution H of $\mathbf{Y}$. Calibration of such predictions refers to the statistical compatibility of the realized values $\mathbf{y}_1, \ldots, \mathbf{y}_n$ with the predictions $H_1, \ldots, H_n$. For probabilistic forecasts F of a univariate quantity $Y \in \mathbb{R}$, various types of calibration have been established [4]. In particular, a forecast is probabilistically calibrated if its probability integral transform (PIT) is uniformly distributed. The PIT Z_F is given by

$$Z_F = F(Y-) + V(F(Y) - F(Y-)),$$

where V is an independent standard uniform random variable, and the forecast F is given in terms of its cumulative distribution function (CDF). Empirical checks for the uniformity of histograms of PIT values have formed a cornerstone of density forecast evaluation [1, 2, 3].

In this work, we provided notions of calibration for predictions of multivariate outcomes $\mathbf{Y} \in \mathbb{R}^d$, that naturally generalize the univariate quantities. In the paper [7], we also proposed tools for empirical calibration checks and illustrated them in various simulation studies. They were applied to compare raw numerical model and statistically postprocessed ensemble forecasts of bivariate wind vectors.

We work in the prediction space setting introduced by [4]. Let $(\Omega, \mathscr{A}, \mathbb{Q})$ be some probability space, $\mathbf{Y} : \Omega \to \mathbb{R}^d$ a random vector, and H a CDF-valued random quantity that is measurable with respect to some sub-σ-algebra $\mathscr{A}_0 \subset \mathscr{A}$, which is understood to encode the information of the forecaster predicting the distribution H. Furthermore, let V be a standard uniform random variable which is independent of $\mathbf{Y}$ and $\mathscr{A}_0$. The predictive distribution H is called *ideal* with respect to $\mathscr{A}_0$, if $H = \mathscr{L}(\mathbf{Y} \mid \mathscr{A}_0)$, that is, H is the conditional law of $\mathbf{Y}$ given $\mathscr{A}_0$.

In the prediction space setting, we define the copula probability integral transform (CopPIT) of H as

$$U_H = \mathscr{K}_H\{H(\mathbf{Y})-\} + V[\mathscr{K}_H\{H(\mathbf{Y})\} - \mathscr{K}_H\{H(\mathbf{Y})-\}],$$

where $\mathscr{K}_H$, denotes the Kendall distribution of the CDF H, that is, the CDF of the random variable $H(X)$ where $X \sim H$. In the univariate case, a forecast F in the prediction space setting is called *probabilistically calibrated* if Z_F is standard uniformly distributed. We generalize this notion to higher dimensions by calling H *probabilistically copula calibrated* if its CopPIT U_H is standard uniformly distributed. If the marginal distributions of H are probabilistically calibrated then probabilistic copula calibration of H only depends on the copula of H, which explains the name. Probabilistic copula calibration can be assessed empirically by considering the histogram of realized CopPIT values.

The CopPIT histogram generalizes the multivariate rank histogram for ensemble forecasts which as introduced in [5] to general probabilistic forecasts, hence, in particular also to density forecasts. Ensemble forecasts correspond to the situation where the probabilistic forecast H is an empirical CDF usually supported in a small number of points $\{\mathbf{x}_1, \ldots, \mathbf{x}_{m-1}\}$. Ensemble forecasts are particularly relevant in weather and climate predictions.

The authors of [6] propose to generalize the multivariate rank histogram of [5] considering the following general procedure. Let $S = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ be an ensemble with verifying observation $\mathbf{x}_m = \mathbf{y}$.

(1) Calculate a univariate pre-rank $\rho(\mathbf{x})$ for each $\mathbf{x} \in S$.
(2) The rank R of $\mathbf{y}$ is the rank of $\rho(\mathbf{x}_m) = \rho(\mathbf{y})$ in $\{\rho(\mathbf{x}_1), \ldots, \rho(\mathbf{x}_m)\}$ with ties resolved at random.

If the elements of S are realizations of iid random variables, then R is uniformly distributed on $\{1, \ldots, m\}$. Examples of possible pre-rank functions are

$$\rho(\mathbf{x}) = \sum_{i=1}^{m} \prod_{k=1}^{d} \mathbf{1}\{\mathrm{rank}(x_{ik}) \leq \mathrm{rank}(x_k)\}$$

for the multivariate rank histogram of [5] ,

$$\rho(\mathbf{x}) = \frac{1}{d} \sum_{k=1}^{d} \mathrm{rank}(x_k)$$

for the average rank histogram, and

$$\rho(\mathbf{x}) = \frac{1}{d}\sum_{k=1}^{d}\big(m - \operatorname{rank}(x_k)\big)\big(\operatorname{rank}(x_k) - 1\big) + (m-1)$$

for the band depth histogram. The last two examples have been proposed by [6]. Multivariate rank histograms have little power when the dimension d is large relative to the ensemble size m (for example, $d = 20$ and $m = 8$), that is, they are essentially flat no matter how miscalibrated the ensemble is. Perhaps surprisingly, the authors of [6] show that their proposed alternatives, the average rank histogram and the band depth histogram, do not suffer from this problem. However, in their current formulation, they are only applicable to ensemble forecasts.

Analogously to the generalization of the multivariate rank histogram to the CopPIT histogram, the general two-step procedure based on pre-ranks described above, can considered for arbitrary predictive distributions H as follows.

Let $F_1, \dots, F_d$ be the marginal CDFs of H. Define

$$\mathbf{Z}_H(\mathbf{Y}) := (Z_{F_1}, \dots, Z_{F_d}),$$

where Z_{F_i} is the univariate PIT of the i-th component of $\mathbf{Y}$. This amounts to taking componentwise ranks (up to scaling) in case of H being an empirical distribution function. Let $\mathfrak{T} : [0,1]^d \to [0,1]$ be a mapping. The quantity $\mathfrak{T}(\mathbf{Z}_H(\mathbf{Y}))$ corresponds to the pre-rank of the observation with respect to the ensemble members. Note that $\mathfrak{T}$ may, or may not depend on the copula C of H.

Let $\mathbf{U} \sim C$, and let $K_C^{\mathfrak{T}}$ be the distribution function of $\mathfrak{T}(\mathbf{U})$. Then, if $\mathbf{X} \sim H$, for some deterministic CDF H, we obtain that $K_C^{\mathfrak{T}}(\mathfrak{T}(\mathbf{Z}_H(\mathbf{X})))$ is standard uniformly distributed, which motivates to define uniformity of $K_C^{\mathfrak{T}}(\mathfrak{T}(\mathbf{Z}_H(\mathbf{Y})))$ as a notion of calibration in the prediction space setting.

For the multivariate rank histogram, we have $\mathfrak{T} = C$, and $K_C^{\mathfrak{T}}$ is the Kendall distribution function of C, which yields the notion of probabilistic copula calibration. For the average rank histogram, $\mathfrak{T}$ is defined as

$$\mathfrak{T} : [0,1]^d \to [0,1],\ u \mapsto \frac{1}{d}\sum_{i=1}^{d} u_i = \frac{\|u\|_{\ell^1}}{d},$$

and for the band depth histogram, we have

$$\mathfrak{T} : [0,1]^d \to [0,1],\ u \mapsto \frac{4}{d}\big(\|u\|_{\ell^1} - \|u\|_{\ell^2}^2\big)$$

if the ensemble size m is large. One would hope that this newly proposed viewpoint on the different notions of calibration for multivariate probabilistic predictions might shed some light on their differences, advantages and disadvantages.

References

[1] A. P. Dawid, *Statistical theory: The prequential approach*, Journal of the Royal Statistical Society: Series A *147* (1984), 278–290.

[2] F. X. Diebold, T. A. Gunther and A. S. Tay, *Evaluating density forecasts with applications to financial risk management*, International Economic Review **39** (1998), 863–883.

[3] T. Gneiting, F. Balabdaoui and A. E. Raftery, *Probabilistic forecasts, calibration and sharpness*, Journal of the Royal Statistical Society: Series B *69* (2007), 243–268.
[4] T. Gneiting and R. Ranjan, *Combining predictive distributions*, Electronic Journal of Statistics **7** (2013), 1747–1782.
[5] T. Gneiting, L. I. Stanberry, E. P. Grimit, L. Held and N. A. Johnson, *Assessing probabilistic forecasts of multivariate quantities, with an application to ensemble predictions of surface winds*, Test **17** (2008), 211-235.
[6] T. Thorarinsdottir, M. Scheuerer and C. Heinz, *Assessing the calibration of high-dimensional ensemble forecasts using the rank histogram*, Journal of Computational and Graphical Statistics, in press (2014), DOI:10.1080/10618600.2014.977477.
[7] J. F. Ziegel and T. Gneiting, *Copula Calibration*, Electronic Journal of Statistics **8** (2014), 2619–2638.

Dynamic asset correlations based on Vines

Jean-David Fermanian
(joint work with Benjamin Poignard)

We propose a new way of generating dynamics of vectors of asset returns, staying inside the family of Multivariate Garch models (see [2]). In such an approach, we specify the first two moments of vectors of returns conditional on their past (and current market information possibly), i.e. we specify their conditional mean and their conditional variance-covariance matrix. Once this is done, some assumed vectors of innovations close the model specification.

Here, a vector of N asset returns r_t is decomposed as $r_t = \mu_t + \epsilon_t$, where μ_t denotes the vector of expected returns conditional on all information at time t, and ϵ_t has a (conditional) zero mean. While the financial literature about asset return predictability tends to focus on the modeling of μ_t, the modeling of time-varying covariances or correlations concentrates on the conditional variance H_t of ϵ_t, which is the "detrended" asset return.

Univariate GARCH dynamics are chosen to get the conditional variance process of (ϵ_t). By stacking these instantaneous conditional variances, we get a sequence of diagonal matrices (D_t). Then, after having specified the ϵ_t-conditional correlation matrices (R_t), we obtain H_t by $H_t = D_t^{1/2} R_t D_t^{1/2}$. This was the method proposed by Engle [5] with the Dynamic Conditional Correlation (DCC). But in a DCC-type model, one has to rely on intricate normalizations to the ϵ_t-correlation matrices. This makes the interpretation of the (R_t) dynamics not intuitive, because it is deduced from another intermediate process, not directly from data.

We develop an alternative methodology which ensures both parsimony and positive definiteness without relying on any normalization. Basically, the idea relies on the modeling of a set of partial correlations, enabling to parameterize any correlation matrix. We choose first a regular vine arbitrarily (see [3]. Then, we obtain an associated set of partial correlations. And a one-to-one mapping between these $N(N-1)/2$ partial correlations and the $N(N-1)/2$ "usual" correlations allows the calculation of R_t. Partial correlations are stacked in a vector Pc_t and the usual

correlations are the coefficients of R_t. We order partial correlations lexicographically, from the shortest to the longest sets of indices. Our proposed correlation dynamics are

$$\Psi\left(Pc_t\right) = \Omega + \Xi\Psi\left(Pc_{t-1}\right) + \Lambda\hat{\zeta}_{t-1}, \quad R_t = vechof\left(F_{vine}\left(\Psi^{-1}(Pc_t)\right)\right),$$

where

- $vechof(\cdot)$ denotes the operator "devectorization", that transforms a vector into a symmetric matrix.
- Ξ and Λ are $N(N-1)/2 \times N(N-1)/2$ squared matrices of unknown parameters, and Ω is an $N(N-1)/2$ unknown vector.
- The vector Pc_t is the "partial correlation vector" deduced from a given R-vine structure.
- The vector $\hat{\zeta}_t$ consists of a relevant function of the t-"innovations".
- Ψ is a function from $]-1,1[^{N(N-1)/2}$ to $\mathbb{R}^{N(N-1)/2}$, defined by

$$\Psi\left(Pc_t\right) = \left(\tan\left(\pi\rho_{1,2,t}/2\right), \cdots, \tan\left(\pi\rho_{N,N-1|L_{N,N-1},t}/2\right)\right)'.$$

- The function F_{vine} is the one-to-one mapping from the vector of partial correlations Pc_t to correlations (in R_t) by using a recursive algorithm (see [10]). By this way, any *arbitrary* process of partial correlations (whose indices are given by any R-vine), here (Pc_t), generates automatically a process of *true correlation* matrices (R_t).

The vector $\hat{\zeta}_t$ has to be specified: for any $L \subset \{1, \ldots, N\}$ and $k \notin L$, define $\upsilon_{k|L,t}$ by

$$\upsilon_{k|L,t} = \frac{\epsilon_{k,t} - \mathbb{E}\left[\epsilon_{k,t}|\epsilon_{L,t}\right]}{\sqrt{h_{k|L,t}}},$$

where $\epsilon_{L,t} = (\epsilon_{i,t})_{i\in L}$, and $\mathbb{E}\left[\epsilon_{k,t}|\epsilon_{L,t}\right]$ corresponds to the orthogonal projection of the variable $\epsilon_{k,t}$ on the space spanned by the vector $\epsilon_{L,t}$. The variance of the "residual" $\epsilon_{k,t} - \mathbb{E}\left[\epsilon_{k,t}|\epsilon_{L,t}\right]$ is denoted by $h_{k|L,t}$. The variables $\upsilon_{k|L,t}$ are not really observable, but we can evaluate $\mathbb{E}\left[\epsilon_{k,t}|\epsilon_{L,t}\right]$ and $h_{k|L,t}$ easily, at least under the conditionally Gaussian assumption. Then, are able to get $\hat{\upsilon}_{k|L,t}$, an approximated value of $\upsilon_{k|L,t}$. Then, the $N(N-1)/2$-sized vector of "innovations" $\hat{\zeta}_t$ stacks the variables $\hat{\upsilon}_{i|L,t}\hat{\upsilon}_{j|L,t}$, when $(i,j|L)$ is an edge of the underlying vine. The order of these edges in ζ_t will be the same as for Pc_t.

We estimate the latter model by a Gaussian Quasi-Maximum Likehood methodology (full estimation and/or sequentially). We evaluate the performances of such specifications by simulation. Such vine-Garch models behave significantly better than scalar DCC-type models, studying a large number of randomly generated 5-dimensional portfolios. Finally, we estimate the asset returns dynamics of several real portfolios based on the MSCI stock indices.

References

[1] Aas, K., Czado, C., Frigessi, A. and Bakken, H., *Pair-copula constructions of multiple dependence*, Working paper 487, Munich University (2006).

[2] Bauwens, L., Laurent, S. and Rombouts, J. , *Multivariate GARCH models: a survey*, Journal of Applied Econometrics **21** (2006), 79–109.

[3] Bedford, T. and Cooke, R., *Vines – a new graphical model for dependent random variables*, Ann. Statist. **30(4)** (2002), 1031–1068.

[4] Dissmann, J., Brechmann, E.C., Czado, C. and Kurowicka, D., *Selecting and estimating regular vine copulae and application to financial returns*, arXiv 1202.2002 (2012).

[5] Engle R.F., *Dynamic conditional correlation: a simple class of multivariate GARCH models*, Journal of Business and Economic Statistics **20** (2002), 339–350.

[6] Engle, R.F. and Sheppard, K., *Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH*, Working Paper No. 8554, National Bureau of Economic Research (2001).

[7] Fermanian, J.-D. and Malongo, H., *Stationarity of dynamic conditional correlation models*, Working Paper Crest 2013-26 (2013).

[8] Kurowicka, D. and Cooke, R., *A parametrization of positive definite matrices in terms of partial correlation vines,* Linear Algebra and its Applications **372** (2003), 225–251.

[9] Kurowicka, D., *Optimal truncation of vines*, in Dependence Modeling: Handbook on Vine Copulae. World Scientific Publishing Co. (2011), 233–248.

[10] Lewandowski, D., Kurowicka, D. and Joe, H., *Generating random correlation matrices based on vines and extended onion method,* Journal of Multivariate Analysis **100** (2009), 1989–2001.

Open problems in risk aggregation under dependence modelling

Valeria Bignozzi

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be an atomless probability space and $L^0 := L^0(\Omega, \mathscr{F}, \mathbb{P})$ be the space of all measurable random variables. A law-invariant risk measure is a functional ρ that assigns to every financial loss X in a given set $\mathscr{X} \subseteq L^0$ a real number (possibly $\pm\infty$), $\rho : X \mapsto \rho(X)$. Such a number, that depends uniquely on the probability distribution of X, is used for regulatory purposes or for internal risk management. The evaluation of $\rho(X)$ is mainly a numerical issue once the distribution of X has been chosen or statistically evaluated. When the distribution is unknown, the risk measure is subject to model uncertainty. This is typically the case with the risk measurement of high dimensional portfolios and has recently gathered a lot of interest in the actuarial and financial literature. We assume that a financial institution holds a d-dimensional risk portfolio, represented by a random vector $\boldsymbol{X} = (X_1, \dots, X_d)$, $X_i \in \mathscr{X}$, $1 \leq i \leq d$. The total loss exposure associated with $\boldsymbol{X}$ is given by the sum

$$X_d^+ = X_1 + \cdots + X_d.$$

Using a risk measure ρ, the aggregate random position X_d^+ is mapped into the real value $\rho(X_d^+)$. Estimating the multivariate distribution for $\boldsymbol{X}$ is a challenging task which is usually performed in two steps: first, d individual models F_i for the marginal loss exposures X_i are independently developed. Then, the marginal distributions are merged into a joint distribution using a dependence structure.

It is in general reasonable to assume that the marginal distributions $F_1, \ldots, F_d$ are known, while $F_{\boldsymbol{X}}$, the joint distribution of $\boldsymbol{X}$, varies in $\mathscr{F}_d(F_1, \ldots, F_d)$, the *Fréchet* class of all possible joint distributions having the fixed marginal models $F_1, \ldots, F_d$. The choice of a single distribution in $\mathscr{F}_d(F_1, \ldots, F_d)$ can lead to the miscalculation of the reserve $\rho(X_d^+)$. The implied model risk is referred to as *dependence uncertainty*.

A natural way to measure dependence uncertainty consists in finding the minimum and maximum possible values of the risk measure ρ evaluated over the class of candidate models. In our framework, we define the smallest and largest capitals to be

$$\underline{\rho}(X_d^+) = \inf\left\{\rho(X_d^+); F_{\boldsymbol{X}} \in \mathscr{F}_d(F_1, \ldots, F_d)\right\}, \tag{1}$$

and

$$\overline{\rho}(X_d^+) = \sup\left\{\rho(X_d^+); F_{\boldsymbol{X}} \in \mathscr{F}_d(F_1, \ldots, F_d)\right\}. \tag{2}$$

Computing the bounds in (1) and (2) if often a difficult task and only partial analytical results are available. There is a wide recent literature that has focused on the computations of these bounds, we refer the interested reader to [4, 1, 6, 9, 3] among many others and to [5] for an overview of the recent developments in this field.

For risk measures that satisfies the subadditivity property, that is

$$\rho(X + Y) \leq \rho(X) + \rho(Y) \quad \text{for any } X, Y \in \mathscr{X}$$

a natural upper bound for a portfolio of risks is given by:

$$\rho(X_d^+) \leq \sum_{i=1}^{d} \rho(X_i).$$

Hence a possible measure of model uncertainty is the *worst superadditivity ratio*

$$\Delta_d = \sup\left\{\frac{\rho(X_d^+)}{\rho(X_1) + \ldots + \rho(X_d)}; F_{\boldsymbol{X}} \in \mathscr{F}_d(F_1, \ldots, F_d)\right\}. \tag{1}$$

We investigate the worst superadditivity ratio in (1) in the context of homogeneous portfolios, $F_1 = \ldots = F_d = F$. Since this ratio depends on the size of the portfolio d, we introduce the notion of *the extreme superadditivity ratio*

$$\Delta^F = \sup_{d \in \mathbb{N}} \frac{\sup\left\{\rho(X_d^+); X_i \sim F, 1 \leq i \leq d\right\}}{n\rho(X_1)} = \frac{1}{\rho(X_1)} \sup_{d \in \mathbb{N}}\left\{\frac{\overline{\rho}(X_d^+)}{n}\right\}.$$

From a mathematical perspective it suffices to focus on the quantity $\Gamma_\rho(F) := \sup_{d \in \mathbb{N}}\left\{\frac{\overline{\rho}(X_d^+)}{n}\right\}$. We show that Γ_ρ is a law-invariant risk measure itself which inherits several interesting properties from ρ. For the class of distortion risk measures, Γ_ρ corresponds to the smallest *coherent* distortion risk measure that dominates ρ.

The results presented in this talk as well as many progresses in the field of dependence modelling are partially based on the recently introduced notions of *complete* and *joint* mixability of a set of distribution functions.

Definition 1 ([11])**.** *The distribution functions $F_1, \dots, F_d$ are said to be* jointly mixable *if there exist d random variables $X_1, \dots, X_d$ such that $X_i \sim F_i, 1 \le i \le d$, and*

$$P(X_1 + \cdots + X_d = C) = 1, \tag{2}$$

for some $C \in \mathbb{R}$. Any such C is called a center *of the jointly mixable distributions and any random vector $(X_1, \dots, X_d)$ satisfying* (2) *is called a* joint mix*. If* (2) *holds with $F_i = F$, $1 \le i \le d$, the distribution F is said to be d-*completely mixable *and the random vector $(X_1, \dots, X_d)$ a* complete mix*.*

The notions of complete and joint mixability are related to some questions of interest in quantitative risk management as for instance: the existence of a least element with respect to the convex order within the set consisting of all sums of random variables with given marginal distributions $F_1, \dots, F_d$; the computation of bounds on the expected value of a supermodular function; the computation of upper and lower sharp bounds on the VaR of a sum of random variables. In the paper *Studying mixability with supermodular aggregating functions* [2] we extend these concepts to supermodular aggregating functions different from the sum.

Definition 2. *Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be a measurable function. The distribution functions $F_1, \dots, F_d$ are said to be* ϕ-jointly mixable *if there exist d random variables $X_1, \dots, X_d$ such that $X_i \sim F_i, 1 \le i \le d$, and*

$$P(\phi(X_1, \dots, X_d) = C) = 1, \tag{3}$$

for some $C \in \mathbb{R}$. Any such C is called a center *of the ϕ-jointly mixable distributions and any random vector $(X_1, \dots, X_d)$ satisfying* (3) *is called a* joint ϕ-mix*. If* (3) *holds with $F_i = F$, $1 \le i \le d$, the distribution F is said to be* ϕ-completely mixable *with index d and the random vector $(X_1, \dots, X_d)$ a* complete ϕ-mix*.*

The concepts of complete/joint/ϕ-joint mixability still present several interesting open questions. We list here some of them, a more detailed treatment of these open problems can be found in [7] and [8]:

(1) Let F be an n-completely mixable distribution, what is a possible copula for an n-complete mix with margins F?
(2) The center of a joint mix is known to be unique under certain assumptions, for instance for bounded distributions. For distributions that are unbounded and with infinite mean, is the center unique?
(3) Proofs on the sufficient conditions for a set of distributions to be mixable are generally long, based on combinatorics and discretizations of the distributions. Is there a more elegant (analytic) proof?

References

[1] Bernard, C., L. Rüschendorf, and S. Vanduffel (2013). *Value-at-Risk bounds with variance constraints.* Preprint, University of Waterloo.
[2] Bignozzi, V. and G. Puccetti (2015). Studying mixability with supermodular aggregating functions. *Statistics & Probability Letters* **100**, 48–55.

[3] Bignozzi, V., G. Puccetti, and L. Rüschendorf (2015). Reducing model risk via positive and negative dependence assumptions. *Insurance Math. Econom.* **61**, 17–26.
[4] Embrechts, P., G. Puccetti, and L. Rüschendorf (2013). *Model uncertainty and VaR aggregation.* J. Bank. Financ. **37(8)**, 2750–2764.
[5] P. Embrechts, G. Puccetti, L. Rüschendorf, R. Wang, A. Beleraj (2014). *An academic response to Basel 3.5.* Risks, **2(1)**, 25–48.
[6] Puccetti, G., B. Wang, and R. Wang (2013). *Complete mixability and asymptotic equivalence of worst-possible VaR and ES estimates.* Insurance Math. Econom. **53(3)**, 821–828.
[7] G. Puccetti and R. Wang. *General extremal dependence concepts.* Preprint, University of Florence, 2015.
[8] R. Wang. *Current open questions in complete mixability.* Preprint, University of Waterloo, 2015.
[9] B. Wang and R. Wang (2015). *Extreme negative dependence and risk aggregation.* J. Multivariate Anal. **136**, 12–25.
[10] Wang, R., V. Bignozzi, and A. Tsanakas (2014). How superadditive can a risk measure be? *Preprint, available at http://ssrn.com/abstract=2373149.*
[11] Wang, R., L. Peng, and J. Yang (2013). Bounds for the sum of dependent risks and worst value-at-risk with monotone marginal densities. *Finance Stoch.* **17(2)**, 395–417.

Extreme value copula estimation based on block maxima of a multivariate stationary time series

Johan Segers
(joint work with Axel Bücher)

The core of the classical block maxima method consists of fitting an extreme value distribution to a sample of maxima over blocks extracted from an underlying series. In asymptotic theory, it is usually postulated that the block maxima are an independent random sample of an extreme value distribution. In practice however, block sizes are finite, so that the extreme value postulate will only hold approximately. A more accurate asymptotic framework is that of a triangular array of block maxima, the block size depending on the size of the underlying sample in such a way that both the block size and the number of blocks within that sample tend to infinity.

The copula of the vector of componentwise maxima in a block is assumed to converge to a limit, which, under mild conditions, is then necessarily an extreme value copula, as proved in [2] and [3]. Under this setting and for absolutely regular stationary sequences, the empirical copula of the sample of vectors of block maxima is shown in [4] to be a consistent and asymptotically normal estimator for the limiting extreme value copula. Moreover, the empirical copula serves as a basis for rank-based, nonparametric estimation of the Pickands dependence function of the extreme value copula, see for instance [1].

The limit distributions of the estimators are the same as if the block maxima had been sampled independently from a distribution whose copula is the limit extreme value copula. The impact of serial dependence is felt rather in the limit copula, which may be different from the extremal attractor of the copula of the stationary distribution of the underlying time series.

The results are illustrated by theoretical examples and a Monte Carlo simulation study.

References

[1] Berghaus, B., A. Bücher, and H. Dette (2013). Minimum distance estimators of the pickands dependence function and related tests of multivariate extreme-value dependence. *Journal de la Société Française de Statistique 154*, 116–137.

[2] Hsing, T. (1989). Extreme value theory for multivariate stationary sequences. *Journal of Multivariate Analysis 29*(2), 274–291.

[3] Hüsler, J. (1990). Multivariate extreme values in stationary random sequences. *Stochastic Processes and their Applications 35*(1), 99–108.

[4] Bücher, A. and Segers, J. (2014). Extreme value copula estimation based on block maxima of a multivariate stationary time series. *Extremes 17*, 495–528.

Recent developments in statistical inference for conditional copulas and association measures

Irène Gijbels

(joint work with Marek Omelka, Noël Veraverbeke)

When modeling the dependence structure between two random variables, given a covariate, a conditional copula function is a useful quantity. In this paper we briefly discuss recent developments in nonparametric estimation of a conditional copula under various (general and restricted) settings.

1. Conditional copulas

The interest is in revealing how the dependence structure between random variables Y_1 and Y_2, varies with a given value of a covariate X. Denoting the conditional joint distribution function of (Y_1, Y_2) given $X = x$ by

$$H_x(y_1, y_2) = P\{Y_1 \leq y_1, Y_2 \leq y_2 | X = x\}$$

and the conditional marginal distribution functions by

$$F_{1x}(y_1) = P\{Y_1 \leq y_1 | X = x\} \qquad \text{and} \qquad F_{2x}(y_2) = P\{Y_2 \leq y_2 | X = x\}$$

the conditional copula function couples the conditional joint distribution function with the conditional marginals, i.e.

$$H_x(y_1, y_2) = C_x\left(F_{1x}(y_1), F_{2x}(y_2)\right). \tag{1}$$

See [3] and [4], among others. If the conditional distributions F_{1x} and F_{2x} are continuous, then the conditional copula function C_x is unique.

Important is to note that the dependence on the value x in (1) comes in at two levels: on the level of the copula function C_x and on the level of the marginals F_{1x} and F_{2x}. In applications it is often assumed that $C_x = C$, i.e. that the dependence on x only comes in through the marginals. Obviously, wrongly assuming this assumption can lead to wrong conclusions. On the other hand, if the assumption

holds, a more efficient estimation should be possible, exploiting this knowledge. This assumption, called the simplifying assumption, leads to the setting

$$H_x(y_1, y_2) = C\left(F_{1x}(y_1), F_{2x}(y_2)\right). \tag{2}$$

Let $(Y_{11}, Y_{21}, X_1), \ldots, (Y_{1n}, Y_{2n}, X_n)$ be an i.i.d. sample of size n from the random vector (Y_1, Y_2, X). The interest is then in nonparametric estimation of the copula function C_x in the general setting, and the function C in the restricted setting. Nonparametric estimators for the conditional copula function C_x, and conditional association measures, have been studied in [5] and [2].

In this short note we discuss nonparametric estimation of the copula function C in the simplified setting (2). We highlight the differences with the nonparametric estimator for C_x, and briefly comment on some asymptotic properties of nonparametric estimators for C in (2). A detailed study is to be found in [1].

2. Nonparametric estimation under the simplifying assumption

Denote by $U_j = F_{jX}(Y_j)$, for $j = 1, 2$. Then, under (2),

$$P\{U_1 \le u_1, U_2 \le u_2 | X\} = P\{Y_1 \le F_{1X}^{-1}(u_1), Y_2 \le F_{2X}^{-1}(u_2) | X\} = C(u_1, u_2).$$

If one would know F_{1x} and F_{2x} one would consider the 'observations'

$$U_{1i} = F_{1X_i}(Y_{1i}) \qquad \text{and} \qquad U_{2i} = F_{2X_i}(Y_{2i}), \qquad i = 1, \ldots, n,$$

and simply build the empirical version of the above probability

$$G_n(u_1, u_2) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\{U_{1i} \le u_1, U_{2i} \le u_2\}$$

resulting in an estimator for the function C

$$C_n^{(\mathrm{or})}(u_1, u_2) = G_n\left(G_{1n}^{-1}(u_1), G_{2n}^{-1}(u_2)\right),$$

where G_{jn}, for $j = 1, 2$ are the marginal distributions obtained from G_n.

If F_{1x} and F_{2x} are unknown, one first obtains estimators $\widehat{F}_{1x}$ and $\widehat{F}_{2x}$. From these the 'pseudo-observations'

$$\widetilde{U}_{1i} = \widehat{F}_{1X_i}(Y_{1i}) \qquad \text{and} \qquad \widetilde{U}_{2i} = \widehat{F}_{2X_i}(Y_{2i}), \qquad i = 1, \ldots, n,$$

are obtained and instead of starting from $G_n(u_1, u_2)$ one uses

$$\widetilde{G}_n(u_1, u_2) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\{\widetilde{U}_{1i} \le u_1, \widetilde{U}_{2i} \le u_2\} \tag{3}$$

leading to a nonparametric estimator for C:

$$C_n(u_1, u_2) = \widetilde{G}_n\left(\widetilde{G}_{1n}^{-1}(u_1), \widetilde{G}_{2n}^{-1}(u_2)\right).$$

This estimator is studied in detail in [1].

The main difference with the nonparametric estimator for C_x (in the general setting), is that in that case the conditional joint distribution depends on x, requiring an extra smoothing operation in the X-domain, leading to working with

$$\widetilde{G}_{xh}(u_1,u_2)=\sum_{i=1}^{n} w_{ni}(x,h_n)\mathbb{I}\{\widetilde{U}_{1i}\le u_1,\widetilde{U}_{2i}\le u_2\}$$

where the weights $w_{ni}(\cdot,h_n)$ involve a bandwidth parametric $h=h_n>0$. Instead of the constant weight $1/n$ in (3), in the simplified setting, one now has weights $w_{ni}(x,h_n)$ appearing in the sum.

A crucial issue in the above is how to estimate the marginal distributions F_{1x} and F_{2x} and which impact this has on the asymptotic behaviour of the final estimator. Various estimation approaches are possible. A very interesting setting is that of assuming location-scale models for Y_1 and Y_2:

$$Y_1=m_1(X)+\sigma_1(X)\,\varepsilon_1,\qquad\qquad Y_2=m_2(X)+\sigma_2(X)\,\varepsilon_2$$

where $m_1,m_2,\sigma_1,\sigma_2$ are either known up to some unknown parameters (i.e. are parametrically specified functions), or fully unknown. The error terms ε_1 and ε_2 are independent of X, with unknown distribution functions $F_{1\varepsilon}$ and $F_{2\varepsilon}$.

Under such general location-scale models: for $j=1,2$

$$F_{jx}(y_j)=F_{j\varepsilon}\left(\tfrac{y_j-m_j(x)}{\sigma_j(x)}\right)$$

and

$$U_{ji}=F_{jX_i}(Y_{ji})=F_{j\varepsilon}\left(\tfrac{Y_{ji}-m_j(X_i)}{\sigma_j(X_i)}\right).$$

Given estimates $\widehat{m}_1$, $\widehat{m}_2$, $\widehat{\sigma}_1$ and $\widehat{\sigma}_2$ of m_1, m_2, σ_1 and σ_2 one then obtains

$$\widehat{F}_{jx}(z)=\widehat{F}_{j\widehat{\varepsilon}}\big(\tfrac{z-\widehat{m}_j(x)}{\widehat{\sigma}_j(x)}\big)$$

where

$$\widehat{F}_{j\widehat{\varepsilon}}(z)=\frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\big\{\widehat{\varepsilon}_{ji}\le z\big\},\qquad\text{with}\qquad\widehat{\varepsilon}_{ji}=\tfrac{Y_{ji}-\widehat{m}_j(X_i)}{\widehat{\sigma}_j(X_i)}$$

and subsequently the estimator $\widetilde{C}_n$ is given by

$$\widetilde{C}_n(u_1,u_2)=\frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\left\{\widehat{F}_{1\widehat{\varepsilon}}(\widehat{\varepsilon}_{1i})\le u_1,\widehat{F}_{2\widehat{\varepsilon}}(\widehat{\varepsilon}_{2i})\le u_2\right\}.$$

When parametric or nonparametric location-scale models for the marginal distributions can be assumed, the discussed estimators perform as well as if the marginals are entirely known. More precisely, the resulting estimators have the same (first-order) asymptotic behaviour as the oracle estimator (case of known marginals). If however the assumption of the parametric/nonparametric location-scale models does not hold (case of misspecification) then these estimators are not consistent.

If one has no idea about appropriateness of location-scale models for the marginals then another, fully nonparametric approach, is to estimate F_{1x} and F_{2x} nonparametrically, for example via

$$\widehat{F}_{1xg_1}(y) = \sum_{k=1}^{n} w_{nk}(x, g_{1n})\, \mathbb{I}\{Y_{1k} \le y\} \qquad \widehat{F}_{2xg_2}(y) = \sum_{k=1}^{n} w_{nk}(x, g_{2n})\, \mathbb{I}\{Y_{2k} \le y\}$$

where additional bandwidth sequences $\{g_{1n}\}$ and $\{g_{2n}\}$ are needed. In contrast to the above approach of a location-scale marginal setting, this estimator is the only (guaranteed) consistent estimator in the simplified setting.

An important issue of further research is to develop nonparametric tests for testing for the simplifying assumption.

Acknowledgement

This research was supported by the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy). The author gratefully acknowledges support from the GOA/12/014 project of the Research Fund KU Leuven.

References

[1] I. Gijbels, M. Omelka, and N. Veraverbeke, *Estimation of a copula when a covariate affects only marginal distributions*, Scandinavian Journal of Statistics, to appear. DOI: 10.1111/sjos.12154

[2] I. Gijbels, N. Veraverbeke, and M. Omelka, *Conditional copulas, association measures and their applications*, Computational Statistics & Data Analysis, **55** (2011), 1919–1932.

[3] R.B. Nelsen, *An Introduction to Copulas*, Second Edition, (2006), Springer, New York.

[4] J. A. Patton, *Modeling asymmetric exchange rate dependence*, International Economic Review, **47** (2006), 527–556.

[5] N. Veraverbeke, M. Omelka, and I. Gijbels, *Estimation of a conditional copula and association measures*, The Scandinavian Journal of Statistic, **38** (2011), 766–780.

Weak convergence of the empirical copula process with respect to weighted metrics

Axel Bücher

(joint work with Betina Berghaus, Stanislav Volgushev)

The empirical copula process plays a central role in the asymptotic analysis of many statistical procedures which are based on copulas or ranks. Among other applications, results regarding its weak convergence can be used to develop asymptotic theory for estimators of dependence measures or copula densities, they allow to derive tests for stochastic independence or specific copula structures, or they may serve as a fundamental tool for the analysis of multivariate rank statistics.

In this talk, we present a weak convergence result for the empirical copula process with respect to weighted supremum distances. For classical empirical processes, such results are well known. For example, the standard 2-dimensional empirical process $\mathbb{F}_n(\mathbf{u}) = \sqrt{n}\{F_n(\mathbf{u}) - F(\mathbf{u})\}$ with F having standard uniform

marginals, converges weakly with respect to the metric induced by the weighted supremum norm

$$\|G\|_\omega = \sup_{\mathbf{u}\in[0,1]^2} \left| \frac{G(\mathbf{u})}{\{g(\mathbf{u})\}^\omega} \right|, \quad g(\mathbf{u}) = \big(\min_{j=1}^{2} u_j \big) \wedge \big(1 - \min_{j=1}^{2} u_j \big),$$

$\omega \in (0, 1/2)$; see, e.g., [2]. For the rank-based empirical copula process $\mathbb{C}_n = \sqrt{n}(\hat{C}_n - C)$, with the empirical copula $\hat{C}_n$, we present a uniform approximation and a weak convergence result with respect to the metric induced by the weighted supremum norm

$$\|G\|_\omega = \sup_{\mathbf{u}\in[0,1]^2} \left| \frac{G(\mathbf{u})}{\{\tilde{g}(\mathbf{u})\}^\omega} \right|, \quad \tilde{g}(\mathbf{u}) = \big(\min_{j=1}^{2} u_j \big) \wedge \big(\min_{j=1}^{2} \{1 - u_j\} \big),$$

for any $\omega \in (0, 1/2)$. The induced distance is stronger than the one based on the function g above.

The usefulness of the result is illustrated by an application to the Pickands estimator for the Pickands dependence function arising in bivariate extreme-value theory. Moreover, extensions of the results to the case of serially dependent time series observations are presented.

All results are taken from [1].

References

[1] B. Berghaus, A. Bücher and S. Volgushev, *Weak convergence of the empirical copula process with respect to weighted metrics*, arXiv:1411.5888 (2014).

[2] C. Genest and J. Segers, *Rank-based inference for bivariate extreme-value copulas*, Ann. Statist. **37(5B)** (2009), 2990–3022.

Conditional copula models with multiple covariates

Elif Acar

Conditional copula models provide a flexible framework to study covariate effects on dependence structures. A number of nonparametric estimation techniques have been recently proposed for these models in the case of a single covariate. See for instance, [3, 6] for a fully nonparametric conditional copula estimation, and [1, 2] for a semiparametric treatment of the problem. These approaches, however, are not directly extendible to, or become impractical in, settings with multiple covariates.

This work proposes a nonparametric modelling strategy that can accommodate multiple covariates in conditional copula models. Let Y_1 and Y_2 denote two continuous response variables, and $X_1, X_2, \ldots X_D$ be the covariates that may affect the strength of dependence between Y_1 and Y_2. The influence of the covariates on the dependence structure can then be modelled by a semiparametric conditional copula model in which the copula is parametric and its parameter varies with covariates. This model is given by

$$(U_1, U_2) \mid X_1, X_2, \ldots, X_D \sim C\{u_1, u_2 \mid \theta(x_1, x_2, \ldots, x_D)\}$$

where $U_k \equiv F_{k|X_1,X_2,\ldots,X_D}(Y_k|x_1, x_2, \ldots, x_D)$ is the conditional marginal distribution of Y_k given $X_1 = x_1, X_2 = x_2, \ldots X_D = x_D$, for $k = 1, 2$, and C is the conditional copula whose parameter depends on the covariate values. Typically, inference for θ involves high-dimensional function estimation, which would be impractical due to the curse of dimensionality. One way to alleviate this issue is to use additive models [4] and formulate the copula parameter function as

$$\theta(X_1, X_2, \ldots X_D) = g^{-1}(\eta(X_1, X_2, \ldots X_D)) = g^{-1}(\eta_0 + \eta_1(X_1) + \ldots + \eta_D(X_D)),$$

where $g^{-1} : \mathbb{R} \to \Theta$ is a pre-specified inverse link function that ensures that the copula parameter has the correct range, and η is the *calibration function* that has an additive structure. The intercept term η_0 is typically included to avoid non-identifiability problems. However, this is not sufficient; the calibration components η_k, $k = 1, \ldots, D$, should be properly centred around zero to guarantee identifiability.

Following [5], we study general identifiability restrictions and empirical ways of centering additive calibration components in conditional copula models. To estimate smooth component functions associated with each covariate, we propose a local likelihood backfitting algorithm, which consists of one-step plug-in estimation followed by mean centering of the resulting estimates. The finite sample performance of the proposed estimator is assessed using simulated data for the cases with two- and five-dimensional covariates. Even with small samples, the backfitting algorithm is able to detect nonlinear/ non constant patterns in the calibration components. We applied the proposed method in a subset of Matched Multiple Births dataset to study the impact of gestational age and mother's age on the dependence between twin birth weights. The results suggest that the strength of dependence changes with both covariates in a nonlinear fashion.

This methodology naturally extends to settings where interest lies in the interdependencies or time-dependent dynamics in multivariate data. One possible extension is via pair copula constructions, which leads to very flexible dependence models for high-dimensional data.

References

[1] Abegaz, F., Gijbels, I. and Veraverbeke, N. [2012] *Semiparametric estimation in copula models*, Journal of Multivariate Analysis, 110, 43–73.

[2] Acar, E. F. , Craiu, R. V. and Yao, F. [2011] *Dependence calibration in conditional copulas: A nonparametric approach*, Biometrics, 67, No 2, 445–453.

[3] Gijbels, I., Veraverbeke, N. and Omelka, M. [2011] *Conditional conditional copulas, association measures and their application*, Computational Statistics Data Analysis, 55, 1919–1932.

[4] Hastie T. and Tibshirani R. [1990] *Generalized Additive Models*, Chapman & Hall / CRC.

[5] Kauerman, G. and Opsomer, J.D. [2003] *Local likelihood estimation in generalized additive models*, Scandinavian Journal of Statistics, **30**, 317–337.

[6] Veraverbeke, N., Omelka, M. and Gijbels, I. [2011] *Estimation of a conditional copula and association measures*, Scandinavian Journal of Statistics, 38, 766–780.

Efficient iterative maximum likelihood estimation

Alexander Ristig
(joint work with Nikolaus Hautsch, Ostap Okhrin)

Statistical inference for models including many parameters is of growing interest in various fields in statistics and econometrics. Such models are mostly estimated using multi-step approaches constructed from parts of the log-likelihood, resulting in inefficient estimators. Prominent examples are the classical two-stage maximum likelihood (ML) estimation procedure according to [3] and the inference functions for margins method to estimate the parameters of multivariate distribution functions, see [2]. In the latter framework, the parameters of the marginal distribution functions are estimated under the assumption of independence in the first stage, whereas the dependence parameter, e.g., a correlation coefficient, is estimated in the second stage.

We address the problem of efficiently estimating a possibly large number of parameters of a complicated log-likelihood function $\ell(\vartheta)$. The ML estimator ϑ_n is typically obtained as solution of numerical optimization procedures as closed-form expressions for ML estimators are rarely available. Pure derivative-based maximization methods require, on the one hand, the first and frequently the second derivative of the log-likelihood function. While analytical derivatives are often difficult to derive because of non-linearities, numerical derivatives are eventually an inaccurate alternative as the approximations might be too rough for non-linear models. In addition to the increasing computational costs based on numerical derivatives, especially if the number of parameters is high, Newton-type optimization procedures suffer from a large number of parameters hampering the inversion of the Hessian. On the other hand, global stochastic optimization routines, like simulated annealing, do not need derivatives, but finding an adequate solution depends on the fine-tuning of the algorithm such as the annealing schedule and acceptance probability. Beyond the fact that convergence can only be obtained, if the algorithm is tuned to be slow, the idea of simulated annealing is not suitable for functions with a lot of arguments, e.g., like non-stochastic hill climbing algorithms. All in all, pure derivative-based as well as pure derivative-free optimization tools might be inappropriate for the one-step maximization of the log-likelihood.

We propose an iterative procedure in the spirit of a Gauß-Seidel algorithm which leads to an asymptotically efficient estimator. The idea is to split the estimation problem into G computationally tractable sub-problems. The approach rests on the assumption of the existence of a consistent but inefficient estimator of a vector of parameters ϑ_n^1. The log-likelihood function is maximized in an iterative procedure where in each sub-step of an iteration, maximization is performed for a sub-vector only while keeping constant the remaining parameters. By iterating through all groups g and successively updating all parameters, we obtain the following algorithm:

Iteration $h > 1$:

$$(g) \quad \vartheta_{g,n}^h = \arg\max_{\vartheta_g} \ell(\vartheta_{1,n}^h, \ldots, \vartheta_{g-1,n}^h, \vartheta_g, \vartheta_{g+1,n}^{h-1}, \ldots, \vartheta_{G,n}^{h-1})$$

The asymptotic distribution of the resulting estimator ϑ_n^h is derived in dependence of the number of underlying iterations h. We moreover show the estimator's asymptotic efficiency as $n \to \infty$, $h \to \infty$ and provide a rule of thumb to approximate the number of iterations h until the procedure comes sufficiently close to the ordinary ML estimator.

Computationally tractable sub-problems are constructed by decomposing the parameter vector ϑ into G sub-vectors $\vartheta_1, \ldots, \vartheta_G$, and $\ell(\cdot)$ is step-by-step maximized with respect to ϑ_g, $g = 1, \ldots, G$, holding fixed all other iteratively updated parameters. The important advantage compared to pure derivative-based and derivative-free optimization techniques is that each sub-maximization of the log-likelihood can be based on an optimization tool specifically suitable for ϑ_g keeping the other sub-vectors fixed, $g = 1, \ldots, G$. For example, first and second derivatives of the log-likelihood function might be difficult to derive for the entire parameter vector ϑ, but relatively easy with respect to some parameters holding the remaining parameters fixed. This combination of different maximization methods makes our procedure tractable even if the underlying model is complex.

Our major focus is on time series models where the number of parameters relative to the number of observations is high (but fixed) and thus it is computationally challenging or virtually impossible to optimize the entire log-likelihood in one step. The proposed algorithm and the corresponding asymptotic theory, however, can also be applied to other estimation and inference problems. We illustrate an application of our procedure and theory to establish the asymptotic properties of the multi-step feasible generalized least squares (FGLS) estimator. Moreover, we show possibilities for efficiently estimating the parameter vector of the DCC model, the multivariate probit model and the stochastic volatility model.

The finite-sample performance of the proposed algorithm is analyzed in two comprehensive simulation studies. The first one investigates the properties of the estimator for a 5-dimensional VARMA model including 17 parameters based on 100 observations. In the second study, we analyze the algorithm's performance to estimate a 15-dimensional Vine-copula containing 105 parameters based on 250 observations. For instance, in the d-dimensional setting of [1], $d(d-1)/2$ stages are required for the Vine-copula. We illustrate that the procedure significantly simplifies the underlying estimation problem and performs sufficiently well.

References

[1] I. Hobæk Haff, *Parameter estimation for pair-copula constructions*, Bernoulli **19(2)** (2013), 462–491.

[2] H. Joe, *Asymptotic efficiency of the two-stage estimation method for copula-based models*, Journal of Multivariate Analysis **94** (2005), 401–419.

[3] H. White, *Estimation, inference and specification analysis*, 1 edition. Cambridge University Press, Cambridge, (1994).

Multivariate Archimax copulas

Anne-Laure Fougères
(joint work with Arthur Charpentier, Christian Genest, Johanna G. Nešlehová)

This talk is a presentation of our joint paper published in 2014 in the Journal of Multivariate Analysis [1].

A d-variate copula is the joint cumulative distribution function of a vector $(U_1, \ldots, U_d)$ of random variables each having a uniform distribution on the interval $(0, 1)$. Following [2], a bivariate copula is said to be Archimax if it can be written, for all $u_1, u_2 \in (0, 1)$, in the form

$$C_{\psi,A}(u_1, u_2) = \psi\left[\{\psi^{-1}(u_1) + \psi^{-1}(u_2)\}A\left\{\frac{\psi^{-1}(u_1)}{\psi^{-1}(u_1) + \psi^{-1}(u_2)}\right\}\right], \tag{1}$$

using maps $A : [0, 1] \to [1/2, 1]$ and $\psi : [0, \infty) \to [0, 1]$ such that

(i) A is convex and, for all $t \in [0, 1]$, $\max(t, 1 - t) \le A(t) \le 1$;
(ii) ψ is convex, decreasing and such that $\psi(0) = 1$ and $\lim_{x\to\infty} \psi(x) = 0$, with the convention that $\psi^{-1}(0) = \inf\{x \ge 0 : \psi(x) = 0\}$.

The term Archimax was chosen by [2] to reflect the fact that if $A \equiv 1$, $C_{\psi,A}$ reduces to an Archimedean copula, viz.

$$C_\psi(u_1, u_2) = \psi\{\psi^{-1}(u_1) + \psi^{-1}(u_2)\}$$

while if $\psi(t) = e^{-t}$ for all $t \in [0, \infty)$, $C_{\psi,A}$ is an extreme-value copula, viz.

$$C_A(u_1, u_2) = \exp\left[\ln(u_1u_2)A\left\{\frac{\ln(u_1)}{\ln(u_1u_2)}\right\}\right].$$

In [2], Archimax copulas are presented as a tool for constructing bivariate distribution functions in the maximum domain of attraction of an extreme-value copula $C_{A^\star}$ where, for all $t \in (0, 1)$,

$$A^\star(t) = \{t^{1/\alpha} + (1 - t)^{1/\alpha}\}^\alpha A^\alpha\left\{\frac{t^{1/\alpha}}{t^{1/\alpha} + (1 - t)^{1/\alpha}}\right\} \tag{2}$$

when the map $t \mapsto \psi^{-1}(1 - 1/t)$ is regularly varying at infinity of degree $-1/\alpha$ with $\alpha \in (0, 1]$; see, e.g., p. 13 in [3] for a definition of regular variation. Bivariate Archimax copulas have been further studied and found to be useful in various contexts since their introduction; see, e.g., [4] for applications in hydrology and `www.math.sk/wiki/bacigal` for a library of R programs.

Recently, [5] and [6] proposed an extension of the family (1) to arbitrary dimension $d \ge 3$. Their generalization involves the notion of stable tail dependence function originally due to [7]. A function $\ell : [0, \infty)^d \to [0, \infty)$ is called a d-variate stable tail dependence function if there exists a d-variate extreme-value copula D such that, for all $x_1, \ldots, x_d \in [0, \infty)$,

$$\ell(x_1, \ldots, x_d) = -\ln\{D(e^{-x_1}, \ldots, e^{-x_d})\}.$$

Let $\psi : [0,\infty) \to [0,1]$ be the generator of a d-variate Archimedean copula C_ψ defined, for all $u_1, \ldots, u_d \in (0,1)$, by

$$C_\psi(u_1, \ldots, u_d) = \psi\{\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)\}.$$

As shown by [8], this occurs if and only if the map $\psi : [0,\infty) \to [0,1]$ satisfies $\psi(0) = 1$, $\lim_{x\to\infty} \psi(x) = 0$ and is d-monotone. The latter property means that ψ has $d-2$ derivatives on $(0,\infty)$ and, for all $j \in \{0, \ldots, d-2\}$, $(-1)^j \psi^{(j)} \geq 0$ with $(-1)^{d-2}\psi^{(d-2)}$ being non-increasing and convex on $(0,\infty)$.

Mesiar & Jágr [6] suggest that a suitable d-variate extension of the notion of Archimax copula would be obtained by setting, for all $u_1, \ldots, u_d \in (0,1)$,

$$C_{\psi,\ell}(u_1, \ldots, u_d) = \psi \circ \ell\{\psi^{-1}(u_1), \ldots, \psi^{-1}(u_d)\}. \tag{3}$$

This is indeed reasonable, as when $d = 2$, one recovers Equation (1) by setting $A(t) = \ell(t, 1-t)$ for all $t \in [0,1]$. While expression (3) appears as formula (18) in [6], these authors merely conjecture that $C_{\psi,\ell}$ is a copula for any choice of ψ and ℓ. This is their Open Problem 4.1.

In joint paper [1], we solve this problem by showing that $C_{\psi,\ell}$ as defined above is indeed a copula for any combination of d-variate Archimedean generator ψ and d-variate stable tail dependence function ℓ. This result is established by combining a composition theorem of [9], a characterization of d-variate Archimedean generators popularized by [8], and a recent characterization of stable tail dependence functions due to [10]. Two different stochastic representations of multivariate Archimax copulas are then provided which shed light on their properties and facilitate simulation; it is also emphasized there that for some d-variate stable tail dependence functions ℓ, the condition on ψ is not necessary for (3) to be a copula. Algorithms for generating observations from this new class of copulas are presented and illustrated. The maximum and minimum attractors of multivariate Archimax copula families are then derived. Finally, a few remaining challenges are outlined, where partial results are offered on the level of dependence that can be achieved by multivariate Archimax copulas.

References

[1] A. Charpentier, A.-L. Fougères, C. Genest, & J. G. Nešlehová, *Multivariate Archimax Copulas*, Journal of Multivariate Analysis **126** (2014), 118–136.

[2] Ph. Capéraà, A.-L. Fougères & C. Genest, *Bivariate distributions with given extreme value attractor*, Journal of Multivariate Analysis **72** (2000), 30–49.

[3] S. I. Resnick, *Extreme Values, Regular Variation and Point Processes*, Springer N.Y. (1987).

[4] T. Bacigál, V. Jágr, & R. Mesiar, *Non-exchangeable random variables, Archimax copulas and their fitting to real data*, Kybernetika **47** (2011), 519–531.

[5] T. Bacigál & R. Mesiar, *3-dimensional Archimax copulas and their fitting to real data*, COMPSTAT 2012 Proceedings (2012).

[6] R. Mesiar & V. Jágr, *d-dimensional dependence functions and Archimax copulas*, Fuzzy Sets and Systems **228** (2013), 78–87.

[7] X. Huang, *Statistics of Bivariate Extreme Values*, Tinbergen Institute Research Series, The Netherlands,(1992).

[8] A. J. McNeil & J. G. Nešlehová, *Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions*, The Annals of Statistics **37** (2009), 3059–3097.

[9] P. M. Morillas, *A characterization of absolutely monotonic* (Δ) *functions of a fixed order*, Publ. Inst. Math. (Beograd) (N.S.) **78(92)** (2005), 93–105.
[10] P. Ressel, *Homogeneous distributions–and a spectral representation of classical mean values and stable tail dependence functions*, Journal of Multivariate Analysis **117** (2013), 246–256.

Copulae based factor model for credit risk analysis

Cathy Yi-Hsuan Chen
(joint work with Meng-Jou Lu, Wolfgang Karl Härdle)

Quantitative methods are important for practical credit risk management. A standard way to access credit risk employs a factor model based on joint multivariate normal distribution properties. A consequent weakness is a thin tail of the risk factors and constant dependence structure. We extend a one-factor Gaussian copula model to make a more accurate forecast of a default probability model. We propose to incorporate state-dependent recovery rate into the conditional factor loading, and model them by sharing a unique common factor. The common factor governs the default rate and recovery rate simultaneously, and creates their association implicitly. Based on Basel III, we also provide the evidence of obligors' behaviors influenced by systematic risk than idiosyncratic risk when period is hectic. Thus, in order to identify the role of the random factor loading and state-dependent recovery rate, we develop four models to evaluate the performance on the default prediction. In addition, we refine the joint Gaussian distribution by a copula based on: The common factor is modeled via a generalized extreme value distribution to determine the performance of forecasting. For robustness testing, we also show the result of sensitivity analysis to the parameters of common factor. Among the models considered, the one with random factor loading and state-dependent recovery rate turns out to be superior.

The main goal behind Basel III is to highlight the importance of systematic risk. First, appropriate capital requirements are important from the perspective of stability of the whole system. Second, [2] states that procyclicality destabilizes the whole systematic risk through amplifying financial shocks. However, since Basel II focused on minimizing the default probability of individuals, this accord failed to guarantee a stable financial system due to a lack of concern for systematic risk. Note that if financial institutions overlook the higher contribution of systematic risk to each firm's value in the bad market situation than idiosyncratic risk, the financial imbalance will be even worse. Especially, systematic risk contributes more to firm's credit risk in a market downturn than in a tranquil market.

Our default's data analysis contains 2008 and 2009, as collected Moody's report. For this period, we employ a state-dependent concept in order to capture an asymmetric impact from the common risk factor. Consequently, we achieve that both conditional factor loading and state-dependent recovery rates improved calibration of our default prediction. In FC model, it overestimates the weight of systematic risk in quiet period. On the other hand, it underestimates the correlation

when market in downturn. We find that incorporation of factor loading into state-dependent recovery rate sharing one unique factor improves the accuracy of default prediction. This result is coherent with goal of Basel III which considers the whole systematic risk rather than the idiosyncratic risk leading to an increase in overall financial stability.

The work presented was:

(1) Conditional Default Model: The systematic latent factor, Z, representing the economic condition that characterizes the systematic credit risk, is of course influencing the default probability $P_i(t)$ and the recovery rate $R_i = 1 - G_i$. So given Z, one may write the conditional default probability $P_i(Z|S = H, Q)$ and conditional LGD, $G_i(Z|S = H, Q)$ as a function of Z. Given the conditional factor loading which proposed by [3], α^H, α^Q, the conditional default model is defined as following,

$$U_i|_{\mathrm{S=H}} = \alpha_i^H Z + \sqrt{1 - (\alpha_i^H)^2}\varepsilon_i$$
$$U_i|_{\mathrm{S=Q}} = \alpha_i^Q Z + \sqrt{1 - (\alpha_i^Q)^2}\varepsilon_i$$

$$P(\tau_i < t|\mathrm{S}) = F\left[\frac{F^{-1}\{P_i(t)\} - \alpha_i^S Z}{\sqrt{1 - (\alpha_i^S)^2}}\right] = P_i(Z|\mathrm{S}) \quad \mathrm{S} \in \{\mathrm{H,Q}\}$$

(2) State-dependent recovery rate: The recovery rate is governed by the state of economy. More specifically, we follow [1] incorporate a conditional correlation structure into this model. Accordingly, we set $R_i(Z|S = H, Q)$, of obligor i, in relation to the common factor Z and the marginal default probability P_i as:

$$G_i(Z|\mathrm{S=H}) = (1 - \bar{R}_i)\frac{F\left[\{F^{-1}\left(\overline{P}_i\right) - \alpha_i^H Z\}/\sqrt{1 - (\alpha_i^H)^2}\right]}{F\left[\{F^{-1}\left(P_i\right) - \alpha_i^H Z\}/\sqrt{1 - (\alpha_i^H)^2}\right]}$$
$$G_i(Z|\mathrm{S=Q}) = (1 - \bar{R}_i)\frac{F\left[\{F^{-1}\left(\overline{P}_i\right) - \alpha_i^Q Z\}/\sqrt{1 - (\alpha_i^Q)^2}\right]}{F\left[\{F^{-1}\left(P_i\right) - \alpha_i^Q Z\}/\sqrt{1 - (\alpha_i^Q)^2}\right]}$$

(3) Based on these specifications, conditional default probability $P_i(Z|\mathrm{S=H,Q})$ and conditional LGD, $G_i(Z|\mathrm{S=H,Q})$, conditional expected loss, therefore, are

$$\mathsf{E}(L_i|Z) = \omega G_i(Z|\mathrm{S=H})P_i(Z|\mathrm{S=H}) + (1 - \omega)G_i(Z|\mathrm{S=Q})P_i(Z|\mathrm{S=Q})$$

References

[1] Amraoui, Salah and Cousot, Laurent and Hitier, Sébastien and Laurent, Jean-Paul, *Pricing CDOs with state-dependent stochastic recovery rates*, Quantitative Finance **12** (2012), 1219–1240.

[2] BCBS, *Strengthening the resilience of the banking sector*, (2009).

[3] Kim, Jongwoo and Finger, Christopher C, *A stress test to incorporate correlation breakdown*, Journal of Risk **2** (2000), 5–20.

Copula based spectral analysis

Holger Dette

(joint work with Marc Hallin, Tobias Kley, Stefan Skowronek, Stanislav Volgushev)

Quantile- and copula-related spectral concepts recently have been considered by various authors. Those spectra, in their most general form, provide a full characterization of the copulas associated with the pairs (X_t, X_{t-k}) in a process $(X_t)_{t\in\mathbb{Z}}$, and account for important dynamic features, such as changes in the conditional shape (skewness, kurtosis), time-irreversibility, or dependence in the extremes, that their traditional counterpart (based on auto-covariances) cannot capture.

In this talk we present results regarding consistency and weak convergence for some of the corresponding estimators proposed in the literature. In particular a detailed asymptotic analysis of a class of smoothed rank-based cross-periodograms associated with the *copula spectral density kernels* is presented. In the second part of the talk the results are extended to relax the assumption of stationarity into much weaker *local stationarity* conditions, which have been developed for a variety of time-series models. In particular we introduce time-varying versions of the copula spectra and periodograms along with a new definition of *strict* local stationarity that allows us to handle completely general non-linear processes without any moment assumptions, thus accommodating our copula based-based concepts and methods. We establish the consistency of our methods, and illustrate their power by means of simulations and an empirical study of the Standard & Poor's 500 series. This empirical study brings evidence of important variations in serial dependence structures both across time (crises and quiet periods exhibit quite different dependence structures) and across quantiles (dependencies between extreme quantiles are not the same as in the "median" range of the series). Such variations remain completely undetected, and are actually undetectable, via classical covariance-based spectral methods.

References

[1] Dette, H., Hallin, M., Kley, T., Volgushev, S. (2015). *Of copulas, quantiles, ranks and spectra: An L_1-approach to spectral analysis.* To appear in: Bernoulli.

[2] Kley, T. Volgushev, S. , Dette, H., Hallin, M., (2013). *Quantile spectral processes – asymptotic analysis and inference.* To appear in: Bernoulli.

[3] Skowronek, S., Volgushev, S., Kley, T., Dette, H., Hallin, M., (2013). *Quantile spectral analysis for locally stationary time series.* Available on Arxiv.

Algebraic properties of the star–product of copulas

Carlo Sempi
(joint work with Francesco Catino)

Denote by $\mathscr{C}_2$ the set of (bivariate) copulas. A copula $C \in \mathscr{C}_2$ is increasing in each variable, viz., for every $s \in \mathbb{I} := [0,1]$ the functions $t \mapsto C(t,s)$ and $t \mapsto C(s,t)$ are increasing, and, hence, differentiable a.e.. Where they exist, the partial derivatives $\partial_1 C(s,t) := \frac{\partial C(s,t)}{\partial s}$ and $\partial_2 C(s,t) := \frac{\partial C(s,t)}{\partial t}$ satisfy

$$0 \le \partial_1 C \le 1\,, \qquad 0 \le \partial_2 C \le 1\,.$$

Darsow et al. [1] have introduced a binary operation on $\mathscr{C}_2$: For every pair of copulas A and B in $\mathscr{C}_2$, define the mapping $* : \mathscr{C}_2 \times \mathscr{C}_2 \to \mathbb{R}_+^{\mathbb{I}^2}$ via

$$(AB)(u,v) = (A * B)(u,v) := \int_0^1 \partial_2 A(u,t)\, \partial_1 B(t,v)\, dt\,.$$

The operation thus defined is called the $*$–product or the *Markov product* of A and B. It has the following properties:

(a) $A * B$ is a copula: $* : \mathscr{C}_2 \times \mathscr{C}_2 \to \mathscr{C}_2$;
(b) the Markov product is associative $A\,(BC) = (AB)\,C$;
(c) $M_2(u,v) := \min\{u,v\}$ is the unit of the product: $CM_2 = M_2C = C$ for every $C \in \mathscr{C}_2$;
(d) $\Pi_2(u,v) := uv$ is the zero of the product: $C\Pi_2 = \Pi_2 C = \Pi_2$ for every $C \in \mathscr{C}_2$.

Thus $(\mathscr{C}_2, *)$ is a monoid (=semigroup with unit), endowed with a zero, but it is not commutative; moreover, the Markov product is both right– and left–distributive over convex combinations. It is our goal to study the algebraic properties of the semigroup $(\mathscr{C}_2, *)$.

A *Markov operator* $T : L^1(\mathbb{I}) \to L^1(\mathbb{I})$ is a linear operator such that

(a) it is positive, $f \ge 0 \Longrightarrow Tf \ge 0$;
(b) it preserves expectations, $\mathbb{E}(Tf) = \mathbb{E}(f)$;
(c) $T\mathbf{1} = \mathbf{1}$.

There exists a one–to–one correspondence between copulas and Markov operators; this is achieved by the equations

$$(T_C f)(x) = \frac{d}{dx} \int_0^1 \partial_2 C(x,t)\, f(t)\, dt\,,$$

$$C_T(u,v) = \int_0^u T\mathbf{1}_{[0,v]}(s)\, ds\,.$$

Moreover, if A and B are copulæ, then, under the correspondence just introduced one has

$$T_A \circ T_B = T_{AB}\,.$$

It follows from Pfanzagl's characterisation [4] of Conditional Expections (CEs) that the following conditions are equivalent for a Markov operator $T : L^1(\mathbb{I}) \to L^1(\mathbb{I})$:

(a) T is the restriction to $L^1(\mathbb{I})$ of a CE on $(\mathbb{I}, \mathscr{B}(\mathbb{I}), \lambda)$;
(b) T is idempotent, viz. $T^2 = T$.

When either condition is satisfied, then $T = \mathbb{E}_{\mathscr{G}}$, where

$$\mathscr{G} := \{A \in \mathscr{B}(\mathbb{I}) : T\,\mathbf{1}_A = \mathbf{1}_A\}\ .$$

A copula C is said to be *idempotent* if $C\,C = C$, viz., if it satisfies the integro–differential equation

$$C(u,v) = \int_0^1 \partial_2 C(u,t)\,\partial_1 C(t,v)\,dt\,.$$

Idempotent copulæ and Markov operators are related.

Theorem 1. [5] *For a copula $C \in \mathscr{C}_2$, the following statements are equivalent:*

(a) *the corresponding Markov operator T_C is a* CE*;*
(b) *the corresponding Markov operator T_C is idempotent;*
(c) *C is idempotent.*

As a consequence to every sub–σ–field $\mathscr{G}$ of $\mathscr{B}$, the Borel σ–field of $\mathbb{I}$, there corresponds a unique idempotent copula $C(\mathscr{G}) \in \mathscr{C}_2$ such that $\mathbb{E}_{\mathscr{G}} = T_{C(\mathscr{G})}$. Conversely, to every idempotent copula $C \in \mathscr{C}_2$ there corresponds a unique sub–σ–field $\mathscr{G}(C)$ of $\mathscr{B}$ such that $T_C = \mathbb{E}_{\mathscr{G}(C)}$.

Let $\mathscr{I}_2$ denote the set of idempotents of $(\mathscr{C}_2, *)$. Every idempotent copula $A \in \mathscr{I}_2$ is symmetric (=exchangeable)[2], $A^T = A$, i.e., $A(u,v) = A(v,u)$ for every (u,v) in $\mathbb{I}^2$. Since the product of two idempotents need not be an idempotent, $(\mathscr{I}_2, *)$ is not a sub–semigroup of $(C_2, *)$. However, if two idempotents $A \in \mathscr{I}_2$ and $B \in \mathscr{I}_2$ commute, viz., $AB = BA$, then also AB is an idempotent, viz., $(AB)^2 = AB$.

A partial order $\leq_*$ is introduced on $\mathscr{C}_2$: $A \leq_* B$ if two copulæ $X \in \mathscr{C}_2$ and $Y \in \mathscr{C}_2$ exist such that

$$\begin{cases} A = XB = BY\,, \\ A = XA\,. \end{cases}$$

It follows from semigroup theory, see [3], that the relation $\leq_*$ on $\mathscr{C}_2$ is a partial order. When restricted to $\mathscr{I}_2$ the partial order $\leq_*$ reads $A \leq_* B$ if, and only if, $A = AB = BA$. It suffices to take $X = Y = A$.

Theorem 2. *Let $\mathscr{G}_1$ and $\mathscr{G}_2$ be sub–σ–fields of $\mathscr{B}$ and let C_1 and C_2 the (uniquely determined) idempotent copulæ corresponding to them. Then the following statements are equivalent:*

(a) $\mathscr{G}_1 \subseteq \mathscr{G}_2$
(b) $C_1 \leq_* C_2$

Recall that if $\mathscr{G}_1 \subseteq \mathscr{G}_2$ and $\mathbb{E}_j := \mathbb{E}_{\mathscr{G}_j}$, then $\mathbb{E}_1\,\mathbb{E}_2 = \mathbb{E}_2\,\mathbb{E}_1 = \mathbb{E}_1$ or $C_1 C_2 = C_2 C_1 = C_1$, viz., $C_1 \leq_* C_2$.

The following result, originally proved in [2], is now immediate.

Corollary *The set $(\mathscr{I}_2, \leq_*)$ of idempotent copulas endowed with the partial order $\leq_*$ is a lattice.*

Proof. Let $C_j \in \mathscr{I}_2 \quad (j = 1, 2)$ and $\mathscr{G}_j$ the corresponding sub–σ–fields of $\mathscr{B}$, let $\mathscr{G} = \mathscr{G}_1 \vee \mathscr{G}_2$ be the σ–field generated by the union $\mathscr{G}_1 \cup \mathscr{G}_2$ and $C \in \mathscr{I}_2$ the unique idempotent copula corresponding to $\mathscr{G}$; then, because of the previous theorem both $C_1 \leq_* C$ and $C_2 \leq_* C$. □

References

[1] W.F. Darsow, B.E. Nguyen and T. Olsen, *Copulas and Markov processes*, Illinois J. Math. **36** (1992), 600–642.

[2] W.F. Darsow and T. Olsen, *Characterization of idempotent copulas*, Note Mat. **30** (2010), 147–177.

[3] H. Mitsch, *A natural partial order for semigroups*, Proc. Amer. Math. Soc. **97** (1986), 384–388.

[4] J. Pfanzagl, *Characterization of conditional expectation*, Ann. Math. Statist. **28** (1967), 415–421.

[5] C. Sempi, *Conditional expectations and idempotent copulæ*, in Distributions with given marginals and statistical modelling, C.M. Cuadras, J. Fortiana, J.A. Rodríguez Lallena, Eds., Kluwer, Dordrecht, (2003) pp. 223–228.

Gaussian copula random field with application to longitudinal neuroimaging data analysis

Peter X. K. Song

(joint work with Jian Kang)

Motivated by the needs of analyzing massive longitudinal imaging data, this paper extends the GeoCopula proposed by [2] and develops a multilevel spatial-temporal copula model for the analysis of high-dimensional imaging data. The resulting model is termed as imageCopula. We propose a two-step composite likelihood approach based on the joint composite estimating functions (JCEF) [1] for parameter estimation and inference. The computation of the proposed algorithm to solve JCEF is scalable to large-scale imaging data. We conduct several simulation studies to evaluate the performance of the proposed models and estimation methods. We apply the imageCopula to analyze a longitudinal PET data set from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study. We beign with the definition of field copula.

Definition 3. *(Field Copula) Let $\{U(\mathbf{s})\}$ be a collection of random variables indexed by location $\mathbf{s} \in \mathscr{A} \subset \mathbb{R}^d$. Suppose that $U(\mathbf{s}) \sim$ Uniform$(0,1)$ for any $\mathbf{s}$. Let $\sigma(\mathscr{A})$ be the σ-algebra of $\mathscr{A}$. Define a joint distribution function $\mathrm{C}_B : [0,1]^{|B|} \to [0,1]$ for any $B \in \sigma(\mathscr{A})$ as follows. If for any $n \geq 1$ and any $\{s_1, \ldots, s_n\} \subset \mathscr{A}$ the n-dimensional joint distribution is given by the following (conditional) copula:*

$$\Pr\left[U(\mathbf{s}_1) \leq u_1, \ldots, U(\mathbf{s}_n) \leq u_n\right] = \mathrm{C}_{\mathbf{s}_1,\ldots,\mathbf{s}_n}(u_1, \ldots, u_n), (u_1, \ldots, u_n) \in [0,1]^n$$

then C *is a field copula on $\sigma(\mathscr{A})$, denoted by $U(\mathbf{s}) \sim$ FieldCopula$(\mathrm{C}, \sigma(\mathscr{A}))$.*

Definition 4. *(Copula Indeced Random Field) Let $\{\zeta(\mathbf{s})\}$ be a collection of random variables indexed by location $\mathbf{s} \in \mathscr{A} \subset \mathbb{R}^d$. Let $F_{\mathbf{s}}$ be the marginal distribution*

function of $\zeta(\mathbf{s})$ *and* C_B *be a field copula on* $\sigma(\mathscr{A})$*. Then* $\{\zeta(\mathbf{s})\}$ *is distributed according to a copula induced random field by marginals* $F_{\mathbf{s}}$ *and field copule* C_B *on* $\sigma(\mathscr{A})$*, denoted by* $\zeta(\mathbf{s}) \sim \mathrm{FieldCopula}[F_{\mathbf{s}}, \mathrm{C}, \sigma(\mathscr{A})]$ *if and only if any* $n \geq 1$ *and any* $\{s_1, \ldots, s_n\} \subset \mathscr{A}$ *and for all* $(a_1, \ldots, a_n) \in \mathbb{R}^n$*,*

$$\Pr\left[\zeta(\mathbf{s}_1) \leq a_1, \ldots, \zeta(\mathbf{s}_n) \leq a_n\right] = \mathrm{C}_{\mathbf{s}_1,\ldots,\mathbf{s}_n}[F_{\mathbf{s}_1}(a_1), F_{\mathbf{s}_2}(a_2), \ldots, F_{\mathbf{s}_n}(a_n)].$$

Definition 5. *(Gaussian Field Copula) Suppose the field copula*

$$\mathrm{C}_{\mathbf{s}_1,\ldots,\mathbf{s}_n}(u_1, \ldots, u_n) = \Phi_n[\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n); \mathbf{\Sigma}(s_1, \ldots, s_n)]$$

where (i) $\Phi_n(\cdot; \mathbf{\Sigma})$ *is an* n*-dimensional multivariate normal distribution with mean* 0 *and spatially varying correlatoin matrix* $\mathbf{\Sigma}$*; (ii)* $\Phi(\cdot) = \Phi_1(\cdot; 0, 1)$ *is the standard normal distribution function; and (iii)* $\mathbf{\Sigma}(s_1, \ldots, s_n) = \begin{pmatrix} 1 & \ldots & k(\mathbf{s}_1, \mathbf{s}_n) \\ k(\mathbf{s}_2, \mathbf{s}_1) & \ldots & k(\mathbf{s}_2, \mathbf{s}_n) \\ \ldots & \ldots & \ldots \\ k(\mathbf{s}_n, \mathbf{s}_1) & \ldots & 1 \end{pmatrix}$ *with* $k(\mathbf{s}, \mathbf{s}')$ *being a correlation kernel function.*

Moreover, we can define the Gaussian copula induced random field.

Definition 6. *Suppose* $\zeta(\mathbf{s}) \sim \mathrm{FieldCopula}[F_{\mathbf{s}}, \mathrm{C}, \sigma(\mathscr{A})]$*. If* C *is a Gaussian field copula on* $\sigma(\mathscr{A})$ *determined by correlation kernel* $k(\mathbf{s}, \mathbf{s}')$*, then* $\zeta(\mathbf{s})$ *is distributed according to a Gaussian copula induced random field, denoted simply by*

$$\zeta(\mathbf{s}) \sim \mathrm{GaussFieldCopula}[F_{\mathbf{s}}, k]$$

We propose to analyze longitudinal imaging data by a semiparametric location-scale marginal model. For subject $i = 1, \ldots, m$, occasion $t = 1, \ldots, T$ and voxel $v = 1, \ldots, n$,

$$y_i(t, \mathbf{s}_v) = \alpha(t, \mathbf{s}_v) + \sum_{j=1}^{p} x_{i,j}(t)\beta_j(\mathbf{s}_v) + \sum_{k=1}^{q} z_{i,k}\eta_k(\mathbf{s}_v) + \tau(t, \mathbf{s}_v)\zeta_i(t, \mathbf{s}_v),$$

where $\zeta_i(t, \mathbf{s}_v) \overset{\text{iid}}{\sim} \mathrm{GaussFieldCopula}(F_{t,\mathbf{s}_v}, k_t)$, $y_i(t, \mathbf{s}_v) \in \mathbb{R}$ is a longitudinal imaging outcome, $x_{i,j}(t)$ is a vector of time varying covariates, $z_{i,k}$ is a vector of time independent covariates, $\alpha(t, \mathbf{s}_v)$ is a spatially-temporally varying intercept, $\beta_j(\mathbf{s}_v)$ and $\eta_k(\mathbf{s}_v)$ are the vectors of spatially varying coefficients, and $\zeta_i(t, \mathbf{s}_v)$ is a spatio-temporal random field with a correlation kernel function $k(\cdot, \cdot|\theta)$. We choose a non-separable spatiotemporal covariance structure used by [1].

We consider spatial basis expansion on the varying coefficients, and the resulting vector of coefficients is denoted by $\mathbf{\Theta}$. Thus the set of all parameters to be estimated is $\widetilde{\mathbf{\Theta}} = (\mathbf{\Theta}, \theta)$. We propose a two-step estimation procedure to estimate $\widetilde{\mathbf{\Theta}}$. The steps of the estimation is listed below.

Step 1 Construct nonparametric estimation for the marginal standard distribution $\widehat{F}_{t,\mathbf{s}}$ for each time point t and for any location $\mathbf{s} \in \mathscr{A}$.

Step 2 To develop the joint composite estimating function (JCEF) [1] to estimate the varying coefficients $\alpha(t, \mathbf{s})$, $\beta_j(\mathbf{s})$, and $\eta_k(\mathbf{s})$ and parameter θ in the correlation kernel in the model. The JCEF is based on the spatial basis expansions on the varying coefficients.

To construct JCEF, we consider patial, temporal and spatio-temporal composite estimating functions, denoted by, respectively, $\Psi_{n,S}(\widetilde{\mathbf{\Theta}}), \Psi_{n,T}(\widetilde{\mathbf{\Theta}}), \Psi_{n,C}(\widetilde{\mathbf{\Theta}})$. The JCEF is defined as a quadratic inference function of the form: $Q_n(\widetilde{\mathbf{\Theta}}) = \Gamma_n(\widetilde{\mathbf{\Theta}})^\top \mathbf{W}^{-1}\Gamma_n(\widetilde{\mathbf{\Theta}})$, where $\Gamma_n(\widetilde{\mathbf{\Theta}}) = [\Psi_{n,S}^\top(\widetilde{\mathbf{\Theta}}), \Psi_{n,T}^\top(\widetilde{\mathbf{\Theta}}), \Psi_{n,C}^\top(\widetilde{\mathbf{\Theta}})]^\top$ and $\mathbf{W}$ is a positive definite matrix of weights. The estimator is given by $\widehat{\widetilde{\mathbf{\Theta}}} = \arg\min_{\widetilde{\mathbf{\Theta}}} Q_n(\widetilde{\mathbf{\Theta}})$. The estimation of the weight matrix $\mathbf{W}$ (a covariance matrix) can be obtained by a subsampling approach [1]. Based on $\widehat{\widetilde{\mathbf{\Theta}}}$, we can construct estimators for $\alpha(t,\mathbf{s})$, $\beta_j(\mathbf{s})$ and $\eta_k(t,\mathbf{s})$, respectively, given by $\widehat{\alpha}(t,\mathbf{s}) = \sum_{l=1}^{L_n} \widehat{\theta}_{\alpha,t,l}\psi_l(\mathbf{s}), \widehat{\beta}_j(\mathbf{s}) = \sum_{l=1}^{L_n} \widehat{\theta}_{\beta,j,l}\psi_l(\mathbf{s}), \widehat{\eta}_k(\mathbf{s}) = \sum_{l=1}^{L_n} \widehat{\theta}_{\eta,k,l}\psi_l(\mathbf{s})$.

References

[1] Y. Bai, P.X.K. Song and T.E. Raghunathan, *Joint composite estimating functions in spatio-temporal models*, Journal of the Royal Statistical Society Series B **74** (2012), 799–824.

[2] Y. Bai, J. Kang and P.X.K. Song, *Efficient pairwise composite likelihood estimation for spatial-clustered data*, Biometrics **70** (2014), 661–670.

Goodness-of-fit test for specification of semiparametric copula dependence models

Ostap Okhrin

(joint work with Shulin Zhang, Qian M. Zhou, Peter X.-K. Song)

This paper concerns goodness-of-fit tests for semiparametric copula models. Suppose that $X_1 = (X_{11}, \ldots, X_{1d})^T, \ldots, X_n = (X_{n1}, \ldots, X_{nd})^T$ is a random sample of size n drawn from a multivariate distribution $H(x) = H(x_1, x_2, \ldots, x_d)$ with continuous marginal CDF $F(x) \triangleq \{F_1(x_1), \ldots, F_d(x_d)\}$. According to Sklar's theorem [3], we suppose that:

$$H(x_1, x_2, \ldots, x_d) = C_0\{F(x)\} = C_0\{F_1(x_1), \ldots, F_d(x_d)\},$$

where $C_0(\cdot)$ is the true copula function. In practice, we often assume that the underlying true copula C_0 belongs to a parametric class, say, $\mathscr{C} \triangleq \{C(\cdot;\theta), \theta \in \Theta\}$, where $\Theta \subset \mathscr{R}^p$ is a p-dimensional parameter space.

In the following, we are interested in the development of a goodness-of-fit test on the hypotheses

$$\mathscr{H}_0 : C_0 \in \mathscr{C} = \{C(\cdot;\theta) : \theta \in \Theta\} \quad vs. \quad \mathscr{H}_1 : C_0 \notin \mathscr{C} = \{C(\cdot;\theta) : \theta \in \Theta\}.$$

Let $\hat{\theta}$ be the two-step pseudo maximum likelihood estimator of θ given by

$$\hat{\theta} = \arg\max_{\theta\in\Theta} \sum_{t=1}^{n} l\left\{\tilde{F}(X_t);\theta\right\}, \tag{1}$$

where $l\{\tilde{F}(X_t);\theta\} = \log c\{\tilde{F}_1(X_{t1}), \ldots, \tilde{F}_d(X_{td});\theta\}$ and $\tilde{F}(x) = \{\tilde{F}_1(x_1), \ldots, \tilde{F}_d(x_d)\}$ with $\tilde{F}_k(x_k) = \frac{1}{n+1}\sum_{t=1}^{n} I(X_{tk} \le x_k)$, $k = 1, \ldots, d$, and $I(\cdot)$ being the indicator function.

To present our new test, let us randomly divide the original data $\{X_1, \ldots, X_n\}$ into B blocks and denote the b-th block as $X^b = (X_1^b, \ldots, X_{n_b}^b)$, $b = 1, \ldots, B$. For the simplicity of exposition, we assume that all the blocks have equal size, *say*, $n_b \equiv m$, and hence $mB = n$. Similar to the "jackknife" resampling method (e.g. [1]), we can yield a set of delete-one-block PLMEs $\hat{\theta}_{-b}, 1 \leq b \leq B$, according to the following procedure:

$$\hat{\theta}_{-b} = \arg\max_{\theta \in \Theta} \sum_{b' \neq b}^{B} \sum_{i=1}^{m} l\{\tilde{F}(X_i^{b'}); \theta\}, \quad b = 1, \ldots, B. \tag{2}$$

Precisely, we propose a test statistic of the following form:

$$T_n(m) = \sum_{b=1}^{B} \sum_{i=1}^{m} \left[l\{\tilde{F}(X_i^b); \hat{\theta}\} - l\{\tilde{F}(X_i^b); \hat{\theta}_{-b}\} \right]. \tag{3}$$

For $n \to \infty$ and under suitable regularity conditions and assumption, that there exists a parameter value $\theta^* \in \Theta$ such that $\hat{\theta} \to \theta^*$ we can show that

$$T_n(m) \xrightarrow{pr} \mathbb{E}_0 \left[l_\theta^T \{F(X_1); \theta^*\} S(\theta^*)^{-1} l_\theta \{F(X_1); \theta^*\} \right] = \mathrm{tr} \left\{ S(\theta^*)^{-1} V(\theta^*) \right\},$$

where $S(\theta) \triangleq -\mathbb{E}_0 \left[l_{\theta\theta} \{F(X_1); \theta\} \right]$, $V(\theta) \triangleq \mathbb{E}_0 \left[l_\theta \{F(X_1); \theta\} l_\theta^T \{F(X_1); \theta\} \right]$, and $\mathrm{tr}(A)$ denotes the trace of a matrix A. Here also $l_\theta(u; \theta) = \frac{\partial}{\partial \theta} \log c(u; \theta)$, $l_{\theta\theta}(u; \theta) = \frac{\partial^2}{\partial \theta \partial \theta^T} \log c(u; \theta)$, and $\mathbb{E}_0(\cdot)$ represents the expectation under the true copula C_0.

Under regularity assumptions we show, that under the null hypothesis

$$\sqrt{n} \left(R_n - p \right) \xrightarrow{d} N(0, \sigma_R^2), \quad \text{as } n \to \infty,$$

where σ_R^2 is the asymptotic variance. And, that

$$R_n - T_n(m) = o_p(n^{-1/2}).$$

where the asymptotic test statistics $R_n = \mathrm{tr} \left\{ \hat{S}(\hat{\theta})^{-1} \hat{V}(\hat{\theta}) \right\}$. Further on, we considered the local power of the test in the Pitman sense for a constant $\delta > 0$, assuming $C_1\{F(x)\} \geq C_0\{F(x); \theta_0\}$ for all $x \in \mathbb{R}^d$

$$H_{1,n} : P_n^{C_1, \delta}(x) = C_0\{F(x); \theta_0\} + \frac{\delta}{\sqrt{n}} \left[C_1\{F(x)\} - C_0\{F(x); \theta_0\} \right].$$

Under extra regularity assumption, we showed, that under $H_{1,n}$

$$\sqrt{n} \left(R_n - p \right) \xrightarrow{d} \mathrm{N}\{\delta m(c_0, c_1), \sigma_R^2\}$$

where $m(c_0, c_1) = \mathbb{E}_{c_0} \left[W(X_t) g \left\{ F(X_t); \theta_0 \right\} \right]$ and $W(\cdot)$ is a weighting function. In this work we also considered, the goodness-of-fit test for a class of multivariate time series models constructed by a semi-parametric copula with corresponding asymptotic distributional properties.

In most scenarios for goodness-of-fit tests, including the ones for copula models (e.g. [2]), there exists no single dominate optimal test. Consider q test statistics, denoted by $T_n^{(1)}, T_n^{(2)}, \ldots, T_n^{(q)}$, where the subscript n is the sample size. Suppose that all of them have type I error controlled at a given significance level α under a

common null hypothesis. A hybrid test is constructed as follows: Let $p_n^{(i)}$ denote the corresponding p-value obtained from the test statistic $T_n^{(i)}$, $i = 1, \ldots, q$. A hybrid test, denoted by T_n^{hybrid}, will make decision according to a p-value, defined as

$$p_n^{hybrid} = q \times min\{p_n^{(1)}, \ldots, p_n^{(q)}\}.$$

This hybrid test can control both types of errors $P(p_n^{hybrid} \leq \alpha | H_0) \leq \alpha$ and $P(p_n^{hybrid} \leq \alpha | H_1) \geq \max\left\{\beta_n^1(\alpha/q), \ldots, \beta_n^q(\alpha/q)\right\}$

Huge simulation study has been performed in order to support proposed tests in iid as well as time series setup. Empirical examples shows usefulness of the methodology

References

[1] B. Efron, *The jackknife, the bootstrap, and other resampling plans*, Society for Industrial and Applied Mathematics, Philadelphia, PA (1982).

[2] C. Genest, Rémillard, B., and Beaudoin, D. *Goodness-of-fit tests for copulas: A review and a power study*, Insurance: Mathematics and Economics **44** (2009), 199–213.

[3] A. Sklar, *Fonctions de répartition à n dimensions et leurs marges*, Publ. Inst. Stat. Univ. Paris **8** (1959), 229–231.

Volatility clustering, leverage effect, and copulas

Cathy Ning

(joint work with Dinghai Xu)

This research investigates the volatility clustering and leverage effect of asset returns with copulas. Volatility clustering and leverage effect are both well-known stylized features of financial asset returns, which are mainly captured in a GARCH model and its extensions. In this research, we use vine copula models to capture the possible nonlinearity and asymmetry in these features. [3] employ an univariate copula approach in [2] and find nonlinear and asymmetric volatility clustering in asset returns, i.e., clusters of high volatility are much stronger than clusters of low volatilities. Models accommodates asymmetric volatility clustering can significantly improve the out-of-sample forecasts of Value-at-Risk. [4] use pair copula models allowing for non-Gaussian marginal distributions and document a nonlinear leverage effect. This research extends our previous work and captures both features at the same time via the use of M-vine copulas of [1] for higher dimension dependence and the across time univariate dependence. The advantage of using such a model is that it does not impose a linear structure on the volatility clustering such as a GARCH model and it also allows for nonlinear and asymmetric feedback effect and leverage effect. The more complete vine copula approach modeling both features simultaneously is expected to further improve the forecasts of VaR, which is important in risk management.

References

[1] B. K. Beare and J. Seo, *Vine copula specifications for stationary multivariate Markov chains*, Journal of Time Series Analysis **36** (2015), 228–246.

[2] X. Chen and Y. Fan, *Estimation of copula-based semiparametric time series models*, Journal of Econometrics **130** (2006), 307–335.

[3] C. Ning, D. Xu, T. Wirjanto, *Is volatility clustering of asset returns asymmetric?*, Journal of Banking and Finance **52** (2015), 62–76.

[4] C. Ning, D. Xu, T. Wirjanto, *Modeling the leverage effect with copulas and realized volatility*, Finance Research Letters **5** (2008), 221–227.

Über copulæ

Bruno Rémillard

In the first part of the talk, I state two conjectures about the Kendall distribution. The first one is that the Kendall function K determines uniquely any Archimedean copula family for any $d \geq 2$. The conjecture appears in [1] and [2]. The second one was about the uniqueness of the Kendall distribution for the independence copula. During the talk, Fabrizio Durante disproved the claim; it seems that some singular copula would have the same Kendall function as the independence copula.

In the second part of the talk, which is a joint ongoing work with Johan Segers, I stated weak convergence properties for empirical processes constructed from pseudo-observations based on Rosenblatt transforms. This kind of pseudo-observations arise naturally when one studies pair-copula models. Even if the limiting empirical process depends on unknown parameters, one can show that under the hypothesis of independence, the associated empirical copula process will have a distribution free limit. This is an important result for those using vine models. In particular it means that under the null hypothesis of independence for a given pair, Kendall's tau, Spearman's rho and many other dependence measures behave as if we were working with i.i.d. samples.

References

[1] Genest, C., Quessy, J.-F., and Rémillard, B. (2006). Goodness-of-fit procedures for copula models based on the integral probability transformation. *Scandinavian Journal of Statistics*, **33**, 337–366.

[2] Rémillard, B. (2013). *Statistical Methods for Financial Engineering*. Chapman and Hall/CRC Financial Mathematics Series. Taylor & Francis.

Construction and sampling of Archimedean and nested Archimedean Lévy copulas

Marius Hofert
(joint work with Oliver Grothe)

In the copula world, the Archimedean class is well-known and widely used. A large sub-class of Archimedean copulas can be sampled with a stochastic representation due to [3]. The corresponding algorithm is referred to as *Marshall–Olkin Algorithm*. This algorithm is typically substantially faster than the (more generic) *conditional distribution method* for sampling a general copula, as the former does not require to compute and then invert high-order conditional copulas.

The class of nested Archimedean copulas provides an extension of the class of Archimedean copulas in that it allows one to leave the symmetry of the latter and model more flexible partially symmetric dependencies. Sampling nested Archimedean copulas can be achieved in a fast and numerically stable way with a "Marshall–Olkin-type algorithm" based on a stochastic representation derived by [4] and [2].

Lévy copulas (although not distributional copulas) are used to couple Lévy processes and thus to construct (cross-)dependent (vector-valued) Lévy processes. We focus on positive Lévy copulas as general Lévy copulas may be obtained by gluing together 2^d positive Lévy copulas, each referring to one of the 2^d orthants of the d-dimensional Euclidean space; see [5, Theorem 5.3]. In the Lévy world, Archimedean Lévy copulas are also known. They have been sampled with a "conditional method" similar to the conditional distribution method used for sampling copulas. Archimedean Lévy copulas share with Archimedean copulas a similar functional form and their symmetry. The latter implies, however, that any two (or three, or four etc.) marginal Lévy processes jump with dependence structure specified by the same (marginal) Archimedean Lévy copula.

To circumvent this restriction, we introduce the class of (positive) nested Archimedean Lévy copulas. This class allows one to model hierarchical dependencies between Lévy processes. Furthermore, we present a new, Marshall–Olkin-type algorithm for sampling both positive Archimedean and positive nested Archimedean Lévy copulas. Besides not requiring (inverses of) conditional Lévy copulas to be known, this algorithm also does not suffer from an asymmetric bias introduced by the conditional method in the Lévy framework. For more details, see [1] or `demo(NALC)` in the R package `copula`.

References

[1] Grothe, O., Hofert, M., *Construction and sampling of Archimedean and nested Archimedean Lévy copulas*, Journal of Multivariate Analysis, to appear (2014).

[2] Hofert, M., *A stochastic representation and sampling algorithm for nested Archimedean copulas*, Journal of Statistical Computation and Simulation **82** (9) (2012), 1239–1255.

[3] Marshall, A. W., Olkin, I., *Families of multivariate distributions*, Journal of the American Statistical Association **83** (403) (1988), 834–841.

[4] McNeil, A. J., *Sampling nested Archimedean copulas*, Journal of Statistical Computation and Simulation **78** (6) (2008), 567–581.
[5] Tankov, P., *Lévy processes in finance: Inverse problems and dependence modelling*, Ph.D. thesis, Ecole Polytechnique (2004).

Copulas for small area estimation

Louis-Paul Rivest
(joint work with François Verret, Sophie Baillargeon)

This work is concerned with unit level small area estimation. The goal is to predict the mean value of a variable Y, in each small area, using simple random samples drawn in each area and auxiliary variables x known for all the population units. We assume a linear model for the dependent variable Y,

$$Y_{ij} = x_{ij}^\top \beta + \varepsilon_{ij} \quad i = 1, \ldots, m; j = 1, \ldots, n_i, \tag{1}$$

where β is a vector of known regression parameters, i denotes the small area, n_i is the sample size in small area i and j stands for an observed unit in area i. Let k represent an unobserved unit in small area i, the basic regression prediction for Y_{ik} given x_{ik} is $x_{ik}^\top \beta$. By modeling the residual dependency within small areas, one can also construct $\tilde{\varepsilon}_i$, a prediction of the regression error ε_{ik} for unit k, and use $\tilde{Y}_{ik} = x_{ik}^\top \hat{\beta} + \tilde{\varepsilon}_i$ as a predictor (note that subscript k has been dropped since $\hat{\varepsilon}_i$ is the same for all unobserved units in small area i).

The standard model for the residual dependency is the mixed normal model of [1]. It assumes that the vector of regression errors has a multi-normal distribution with mean 0 and an exchangeable covariance matrix, $\sigma^2 \Sigma(\rho, n_i)$, where σ^2 is the known marginal residual variance and $\Sigma(\rho, n)$ is a $n \times n$ correlation matrix where all off diagonal elements are equal to ρ, the known intra cluster correlation. Under this model the best predictor is linear and is given by

$$\tilde{\varepsilon}_i^\ell = n_i \rho \bar{\varepsilon}_{is} / \{1 + (n_i - 1)\rho\},$$

where $\bar{\varepsilon}_{is}$ is the mean of the n_i errors in small area i. This is the BLUP (for best linear unbiased predictor). Thus the predictor for $\bar{Y}_{iU}$, where subscript U means that the mean is evaluated for the N_i units in small area i and N_i, $(N_i > n_i)$, is the size of area i, is

$$\tilde{\bar{Y}}_{iU} = \bar{x}_{iU}^\top \beta + \frac{n_i}{N_i} \varepsilon_{is} \frac{N_- n_i}{N_i} + \tilde{\varepsilon}_i^\ell. \tag{2}$$

The variance of the prediction error, $E\{(\tilde{\bar{Y}}_{iU} - \bar{Y}_{iU})^2\}$ is, when N_i is much larger than n_i, equal to $MSE_i^\ell = \sigma^2 \rho(1 - \rho)/\{1 + (n_i - 1)\rho\}$. This is the standard g_{1i} variance component of the small area mean squared error, see [3].

This works consider a new class of exchangeable models for the regression errors, constructed in terms of $C_{\alpha,n}(u_1, \ldots, u_n)$, an extendible n-dimensional parametric copula model and and arbitrary cumulative distribution function F_e, with a null expectation. The proposed joint distribution of the errors within small area i is $C_{\alpha,n_i}\{F_e(y_{i1} - x_{i1}^\top \beta), \ldots, F_e(y_{in_i} - x_{in_i}^\top \beta)\}$. Suitable candidates for $C_{\alpha,n}$ are

the Archimedean copulas with a single dependency parameter such as Clayton's, Frank's or Gumbel's family, see [2]. The normal and the t copula associated to correlation matrix $\Sigma(\rho, n_i)$ can also be used. For this large class of models the BLUP is given by (2) and MSE_i^ℓ measures its precision, where ρ the intra-cluster correlation depends on both $C_{\alpha,2}$ and F_e.

For the new non-normal models, the best unbiased predictor (BUP) for ε_{ik} is no longer a linear function of $\{\varepsilon_{ij} : j = 1, \ldots, n_i\}$. It can be evaluated as the conditional expectation of ε_{ik} given $\{\varepsilon_{ij} : j = 1, \ldots, n_i\}$,

$$\tilde{\varepsilon}_i = \int_{\Re} z \frac{c_{\alpha,n_i+1}\{F_e(z), F_e(\varepsilon_{i1}), \ldots, F_e(\varepsilon_{in_i})\}}{c_{\alpha,n_i}\{F_e(\varepsilon_{i1}), \ldots, F_e(\varepsilon_{in_i})\}} F_e(dz)$$

where $c_{\alpha,n}$ denotes the copula density. When N_i is large, the variance of the prediction error for $\bar{Y}_{iU}$ constructed with the BUP is $MSE_i = \sigma^2\rho - E\{E(\varepsilon_{ik}|\varepsilon_{ij} : j = 1, \ldots, n_i)^2\}$. First we evaluate the efficiency, MSE_i/MSE_i^ℓ, of the BLUP with respect to the BUP to find out whether considering non-linear predictors is worthwhile. We show that in a model with normal errors and Clayton's copula, the BLUP's efficiency can be as small as 60%. Thus implementing the BUP appears to be useful.

The parameters of the proposed model are β, the vector of regression parameters, α, the dependency parameter, and F_e the centered marginal error distribution. Parameter estimation is this semi-parametric setting is first considered and $\sqrt{m}$ consistent estimators are derived. For β we suggest using the estimator obtained through the algorithm that fits the normal mixed model of [1]. Parameter α is estimated using the generalization of Kendall's tau to clustered data investigated in [4]. Finally F_e is estimated using the empirical cdf of the regression residuals, $\{e_{ij} = Y_{ij} - x_{ij}^\top\hat{\beta}\}$. Then the EBUP, an estimator of the BUP, calculated with the parameters estimates is evaluated. Table 1 presents some simulations comparing the two predictors, EBLUP vs EBUP (the first E stands for empirical). It features $m = 20$ and 40 small areas containing $N_i = 500$ units and samples of size $n_i = 4$ drawn in each one. For all models $\rho = 0.5$ and $\sigma^2 = 2$ so that $MSE_i^\ell = 0.2$. BHF stands for the normal model of [1], Clayton for a model with Clayton's copula and normal margins and Example 3 has a normal copula and log-normal margins. Table 1 presents aggregated results for the m small areas combined. The EBLUP only involves estimators for β and ρ. In the BHF simulation it should be more accurate that the EBUP as the later uses an estimate of F_e. Table 1 reveals that the difference in precision is small, especially for $m = 40$. For Clayton's model the EBUP does well; the efficiencies of the EBLUP are 84% and 78% respectively. In the simulation with log-normal margins the performance of the EBUP is affected by the outliers that occur with this type of data; still it is slightly better than the EBLUP.

References

[1] G. E. Battese, R. M. Harter, and W. A. Fuller, *An error-components models for prediction of county crop areas using survey and satellite data*, Journal of the American Statistical Association, **83** (1988), 28–36.

TABLE 1. Comparisons of the Monte Carlo prediction errors of the EBLUP versus the EBUP.

	EBLUP		EBUP	
Model	Bias	MSE	Bias	MSE
BHF m=20	-.008	.213	-.008	.221
BHF m=40	.003	.206	.003	.210
Clayton m=20	-.000	.209	-.019	.178
Clayton m=40	-.001	.209	-.008	.164
Example 3 m=20	.010	.206	.000	.200
Example 3 m=40	-.001	.205	.008	.200

[2] J.-M. Mai, & M. Scherer (2012), *Simulating Copulas; Stochastic Models, Sampling Algorithms and Applications. Series in Quantitative Finance: Volume 4.* World Scientific Publishing Company.
[3] J. N. K. Rao (2003), *Small Area Estimation.* John Wiley: New York.
[4] H. Romdhani, L.-P. Rivest, and L. Lakhal, *An exchangeable Kendall's tau for clustered data.* Canadian Journal of Statistics, 2014 **42**, 384–403.

Singular copulas

FABRIZIO DURANTE
(joint work with Juan Fernández-Sánchez and Wolfgang Trutschnig)

Copulas are distribution functions of probability measures on $\mathbb{I}^d := [0,1]^d$ ($d \geq 2$) whose univariate marginals are uniformly distributed on $\mathbb{I}$ (see, e.g., [5]). If $\mathscr{C}_d$ denotes the class of d–dimensional copulas, the probability measure μ_C associated with $C \in \mathscr{C}_d$ can be decomposed in the form

$$\mu_C = \mu_C^{ac} + \mu_C^{s},$$

where μ_C^{ac} is absolutely continuous with respect to the Lebesgue measure λ_d, while μ_C^{s} is singular with respect to λ_d.

Now, while absolutely continuous copulas have been largely considered in the literature (especially in view of their practical convenience in applications), there are many reasons to investigate also the set of singular copulas (or, copulas with a singular component). In fact, in many situations, given a functional $g\colon \mathscr{C}_d \to \mathbb{R}$ the supremum $\sup_{C\in\mathscr{C}_d} g(C)$ is attained by non–absolutely continuous copulas. Examples include the copula that gives the worst–case value-at-risk scenario for a two-asset portfolio at a given level α (see, e.g., [9, 14]) and the maximally asymmetric copulas (see, e.g., [7, 12]).

In order to address properly the study of singular copulas we should aim at providing tools to deal with them in a convenient way. A possibility is provided by the representation of copulas in terms of their Markov kernels (see, e.g., [6, 15]). Specifically, suppose that $(X, \mathbf{Y})$ is a random vector in a given probability space

$(\Omega, \mathscr{F}, \mathbb{P})$ distributed according to $C \in \mathscr{C}_d$. Then, for every Borel set $B \subseteq \mathbb{I}^d$, one has

$$\mu_C(B) = \int_{\mathbb{I}} K_C(x, B_x) d\lambda(x),$$

where $B_x := \{\mathbf{y} \in \mathbb{I}^{d-1} : (x, \mathbf{y}) \in B\}$ and the so-called *Markov kernel* K_C satisfies

$$K_C(X(\omega), B) = \mathbb{E}(1_B \circ Y \mid X)(\omega) \qquad \mathbb{P}\text{–a.s.}$$

For every $C \in \mathscr{C}_d$ and for every $\mathbf{y} \in [0,1]^{d-1}$, one has $\partial_1 C(x, \mathbf{y}) = K_C(x, [\mathbf{0}, \mathbf{y}])$ for λ–almost every $x \in [0,1]$. However, for each $x \in [0,1]$, $\mathbf{y} \mapsto K_C(x, [\mathbf{0}, \mathbf{y}])$ is a proper probability distribution function, a fact that often represents an advantage in some proof strategies.

The Markov kernel of a copula $C \in \mathscr{C}_2$ can be decomposed into

$$E \mapsto K_C(x, E) = K_C^a(x, E) + K_C^s(x, E) + K_C^d(x, E)$$

where $K_C^a(x, \cdot)$ is absolutely continuous with respect to λ, $K_C^s(x, \cdot)$ is singular with respect to λ and has no point masses, and $K_C^d(x, \cdot)$ is discrete for every $x \in \mathbb{I}$.

This latter fact has been exploited in [3] in order to provide an alternative viewpoint to a classical problem in optimal coupling (see also [8]).

Theorem 1. *Let F_1 and F_2 be continuous distribution functions. Set $T = F_2 \circ F_1^{-1}$. Let*

$$p = \sup_{(X_1, X_2) \sim C(F_1, F_2)} \mathbb{P}(X_1 = X_2),$$

where the supremum is calculated over all the possible random pairs with the given univariate marginals.

If T is a non singular transformation, then there exists a copula A_T such that the probability mass of the singular component of A_T is concentrated in $\mathrm{Graph}(T)$. *Moreover, $p = \mu_{A_T}(\mathrm{Graph}(T))$.*

In other words, if the random variables X_1 and X_2 have the interpretation of lifetimes, the maximal probability of joint default is obtained when they are coupled by a copula that concentrates the probability mass of the singular component in the graph of a function that depends on the marginals.

The Markov kernel representation of a copula has been also used in the characterization of complete dependence among random variables. Suppose that two random variables U and V are distributed according to $C \in \mathscr{C}_2$. Then C is called *completely dependent* if $V = h(U)$ almost surely or, equivalently, there exists a λ–measure preserving transformation $h \colon \mathbb{I} \to \mathbb{I}$ such that $K_C(x, E) = \mathbf{1}_E(h(x))$ is the Markov kernel of C. In particular, if h is a bijection, then C is called *mutual completely dependent*. The class of complete dependence copulas is a distinguished subset of singular copulas that is dense in $(\mathscr{C}_2, d_\infty)$, where d_∞ denotes the L^∞ distance. Actually, it can be proved that mutually completely dependent copulas (i.e. shuffles of Min) are dense in $(\mathscr{C}_2, d_\infty)$ (see, e.g., [10]). In particular, there are sequences $(C_n)_{n \in \mathbb{N}}$ of shuffles of Min with $d_\infty(C_n, \Pi) \to 0$, a fact that has been investigated several times in the literature (see, e.g., [10, 16].

The latter results strongly depends on the topology we are considering on the space of copulas. In fact, if we assume that $\mathscr{C}_2$ is equipped with the D_1 metric in [15] or with the topology induced by the $\mathfrak{M}$–convergence in [11], the following result holds (for the proofs see [1, 15]).

Theorem 2. *Complete dependence copulas form a closed set in $(\mathscr{C}_2, D_1)$ and well as in $(\mathscr{C}_2, d_{\mathfrak{M}})$.*

A final aspect that we are interested has been the possibility of describing the relative size such special classes of copulas using tools from Baire categories. These investigations have also been motivated by the necessity of clarifying whether classical assumptions in asymptotic statistics apply to a relatively large set (see, e.g., [13]).

We recall that a subset of a metric space (S, d) is *nowhere dense* if its closure has empty interior; while a set $A \subseteq S$ is *meager* or of *first category* in (S, d) if it is expressible as a countable union of nowhere dense sets. Moreover, a set A is called of *second category* if it is not meager and it is called *co-meager* if $S \setminus A$ is meager.

The following results have been proved in [4].

Theorem 3. *Consider the complete metric space $(\mathscr{C}_d, d_\infty)$ as well as $(\mathscr{C}_d, D_1)$.*

- *The set of absolutely continuous copulas is of first category in $(\mathscr{C}_d, d_\infty)$ as well as in $(\mathscr{C}_d, D_1)$.*
- *The class of (purely) singular copulas is co-meager in $(\mathscr{C}_d, d_\infty)$.*

Similar results have been also proved in [2] for several subclasses of exchangeable and associative copulas. It remains an open question whether the class of purely singular copulas is of second category in $(\mathscr{C}_d, D_1)$.

To conclude, many statistical distributions and methods have been introduced in the last years to provide alternatives to absolutely continuous copulas. This report (and the related investigations) suggests that these structures may also arise in a natural way as solutions of special optimization problems and present properties of some mathematical interest.

References

[1] F. Durante and J. Fernández-Sánchez. On the approximation of copulas via shuffles of Min. *Statist. Probab. Lett.*, 82(10):1761–1767, 2012.

[2] F. Durante, J. Fernández-Sánchez, and W. Trutschnig. Baire category results for exchangeable copulas. *Fuzzy Sets and Systems*, in press, 2015.

[3] F. Durante, J. Fernández-Sánchez, and W. Trutschnig. On the singular components of a copula. *J. Appl. Probab.*, in press, 2015.

[4] F. Durante, J. Fernández-Sánchez, and W. Trutschnig. A typical copula is singular. submitted, 2015.

[5] F. Durante and C. Sempi. *Principles of copula theory.* CRC/Chapman & Hall, London, 2015.

[6] J. Fernández-Sánchez and W. Trutschnig. Conditioning-based metrics on the space of multivariate copulas and their interrelation with uniform and levelwise convergence and iterated function systems. *J. Theoret. Probab.*, 2014. doi=10.1007/s10959-014-0541-4.

[7] E. P. Klement and R. Mesiar. How non-symmetric can a copula be? *Comment. Math. Univ. Carolin.*, 47(1):141–148, 2006.
[8] J.-F. Mai and M. Scherer. Simulating from the copula that generates the maximal probability for a joint default under given (inhomogeneous) marginals.
In V.B. Melas, S. Mignani, P. Monari, and L. Salmaso, editors, *Topics from the 7th International Workshop on Statistical Simulation*, Springer Proceedings in Mathematics & Statistics, pages 333–341. Springer, 2014.
[9] G. D. Makarov. Estimates for the distribution function of the sum of two random variables with given marginal distributions. *Teor. Veroyatnost. i Primenen.*, 26(4):815–817, 1981.
[10] P. Mikusiński, H. Sherwood, and M. D. Taylor. Shuffles of Min. *Stochastica*, 13(1):61–74, 1992.
[11] P. Mikusiński and M. D. Taylor. Some approximations of n-copulas. *Metrika*, 72(3):385–414, 2010.
[12] R. B. Nelsen. Extremes of nonexchangeability. *Statist. Papers*, 48(2):329–336, 2007.
[13] H. Putter and W. R. van Zwet. Resampling: consistency of substitution estimators. *Ann. Statist.*, 24(6):2297–2318, 1996.
[14] L. Rüschendorf. Random variables with maximum sums. *Adv. in Appl. Probab.*, 14(3):623–632, 1982.
[15] W. Trutschnig. On a strong metric on the space of copulas and its induced dependence measure. *J. Math. Anal. Appl.*, 384(2):690–705, 2011.
[16] R. A. Vitale. Approximation by mutually completely dependent processes. *J. Approx. Theory*, 66(2):225–228, 1991.

On some new constructions of copulas

Radko Mesiar

Recently, we have introduced several new kinds of constructions of binary copulas. We recall and exemplify some of them. Based on the results of [8, 10], we introduce first the role of ultramodularity in copula constructions, including some examples.

A copula $C : [0,1] \to [0,1]$ is *ultramodular* [9, 10] if and only if for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in [0,1]^2$ satisfying $\mathbf{x} + \mathbf{y} + \mathbf{z} \in [0,1]^2$ we have

$$C(\mathbf{x} + \mathbf{y} + \mathbf{z}) + C(\mathbf{x}) \geq C(\mathbf{x} + \mathbf{y}) + C(\mathbf{x} + \mathbf{z}). \tag{1}$$

Note that ultramodular copulas are just copulas with convex horizontal and vertical sections.

Out of the three basic copulas W, M and product copula Π given by $\Pi(x,y) = x \cdot y$ only W and Π are ultramodular. However, the upper Fréchet–Hoeffding bound M is ultramodular on the upper left triangle

$$\Delta = \{(x,y) \in [0,1]^2 \mid x \leq y\}.$$

Theorem 4 ([10], Theorem 3.1)**.** *Let $C : [0,1]^2 \to [0,1]$ be an Archimedean copula with a two times differentiable additive generator $t : [0,1] \to [0,\infty]$. Then C is ultramodular if and only if t' is constant or $\frac{1}{t'}$ is a convex function.*

As a consequence of Theorem 4.1 in [10] we get:

Theorem 5. *Let $C_1, C_2, D : [0,1]^2 \to [0,1]$ be a copulas and assume that D is ultramodular. Then, for all monotone non–decreasing functions f_1, f_2, g_1, g_2 :*

$[0,1] \to [0,1]$ *with* $D\left(f_1(x), f_2(x)\right) = D\left(g_1(x), g_2(x)\right) = x$ *for all* $x \in [0,1]$*, also the function* $E : [0,1]^2 \to [0,1]$ *given by*

$$E(x,y) = D\left(C_1\left(f_1(x), g_1(y)\right), C_2\left(f_2(x), g_2(y)\right)\right) \tag{2}$$

is a copula.

Example 1. *Here is example of the construction in Theorem 5:*

For each copula C *and all* $\alpha, \beta \in [0,1]$ *the function* $E : [0,1]^2 \to [0,1]$ *given by* $E(x,y) = C\left(x^\alpha, y^\beta\right) \cdot C\left(x^{1-\alpha}, y^{1-\beta}\right)$ *is a copula (this result was obtained independently in* [7]*, see also* [12]*). Putting* $C = W$ *and* $\alpha = \beta = 0.5$*, we obtain the Clayton copula with parameter -0.5 (see* [16]*) given by* $C_{-0.5}(x,y) = \left(\max\left(\sqrt{x} + \sqrt{y} - 1.0\right)\right)^2$.

Next, we will need the Schur concavity of a copula $D : [0,1]^2 \to [0,1]$ on the upper left triangle $\Delta = \left\{(x,y) \in [0,1]^2 | x \le y\right\}$ only, which means that for all $(x,y) \in \Delta$ and for all $\varepsilon > 0$ with $(x+\varepsilon, y-\varepsilon) \in \Delta$ we have

$$D(x,y) \le D(x+\varepsilon, y-\varepsilon).$$

Theorem 6 ([8])**.** *Let* C *be a binary copula and let* D *be a binary copula which is ultramodular and Schur concave on the upper left triangle* Δ*. Then the function* $D(C, C^\star)$ *is a copula, where* $C^\star$ *is the dual copula given by*

$$C^\star(x,y) = x + y - C(x,y).$$

It is remarkable that $D(C, C^\star)$ in Theorem 6 preserves the ultramodularity and the Schur concavity on Δ of the copulas C and D.

Proposition 1. *Let C, D be binary copulas which are ultramodular and Schur concave on the upper left triangle* Δ*. Then also the copula* $D(C, C^\star)$ *is ultramodular and Schur concave on* Δ*.*

It turns out that the ultramodularity of D is a necessary condition if we want $D(C, C^\star)$ to be a copula for each copula C.

Theorem 7. *Let* D *be a binary copula such that for each binary copula* C *the function* $D(C, C^\star)$ *is a copula. Then* D *is ultramodular on the upper left triangle* Δ*.*

Example 2. *The product copula* Π *is both ultramodular and Schur concave, and thus* $\Pi_C : [0,1]^2 \to [0,1]$ *given by* $\Pi_C(x,y) = C(x,y) \cdot (x + y - C(x,y))$ *is a copula for any (bivariate) copula* C*.*

Next, we recall quadratic constructions of copulas [11] and their stochastic interpretation based on results of [4]. Inspired by the fact that the above introduced copula Π_C can be seen as a composite function,

$$\Pi_C(x,y) = P(x,y,C(x,y)),$$

where P is a quadratic polynomial, $P(x,y,z) = z \cdot (x + y - z)$, we were interested [11] in a complete description of all ternary quadratic polynomials P such that, for any bivariate copula C, the function $P_{(C)} : [0,1]^2 \to [0,1]$ given by

$$P_{(C)}(x,y) = P(x,y,C(x,y)) \tag{3}$$

is a copula.

Theorem 8. *For a copula $C : [0,1]^2 \to [0,1]$, let $P_{(C)}$ be a function defined on $[0,1]^2$ by*

$$P_{(C)}(x,y) = ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j, \quad (x,y) \in [0,1]^2,$$

where $z = C(x,y)$ and $a, \ldots, j \in \mathbb{R}$. Then the following are equivalent.

(i) *For any copula C, $P_{(C)}$ is a copula.*

(ii) *$P_{(C)}$ is given by*

$$P_{(C)}(x,y) = cC^2(x,y) + dxy - cxC(x,y) - cyC(x,y) + (1+c-d)C(x,y),$$

with coefficients c, d satisfying conditions

$$0 \le d \le 1, \ 0 \le d - c \le 1.$$

Observe that, up to the case when $c = d = 0$, the only copula invariant under the quadratic construction generated by the polynomal

$$P(x,y,z) = cz^2 + dxy - cxz - cyz + (1+c-d)z$$

by (3), i.e., $P_{(C)} = C$, is the Plackett copula C_α^{Pl} with parameter $\alpha = \frac{d}{d-c}$ if $d \neq c$, $P_{(C)} = M$ if $d = c \neq 0$. If $c = d = 0$ then $P_{(C)} = C$ for any copula C.

Note that considering $c = -1$ and $d = 0$, the corresponding quadratic construction coincides with construction described in Example 2, i.e.,$P_{(C)} = \Pi_C$. This construction has also an interesting stochastic interpretation, see [4].

Consider IID random vectors (X_1, Y_1) and (X_2, Y_2) characterized by a copula C, and such that X_1, Y_1, X_2, Y_2 are uniformly distributed over $[0,1]$. Then the random vector (Z_1, Z_2) is characterized by the copula Π_C, where

$$(Z_1, Z_2) = \begin{cases} (\min(X_1, X_2), \max(Y_1, Y_2)) & \text{with pp. } 0.5, \\ (\max(X_1, X_2), \min(Y_1, Y_2)) & \text{with pp. } 0.5. \end{cases}$$

Another kind of constructions of singular copulas is related to modular functions [2].

Recall that a function $A : [0,1]^2 \to [0,1]$ is called an aggregation function if it is monotone and satisfies two boundary conditions $A(0,0) = 0$ and $A(1,1) = 1$ [5]. An aggregation function A is said to be modular (1–Lipschitz) if it satisfies, for all $\mathbf{x}, \mathbf{y} \in [0,1]^2$,

$$A(\mathbf{x}) + A(\mathbf{y}) = A(\mathbf{x} \vee \mathbf{y}) + A(\mathbf{x} \wedge \mathbf{y})$$
$$(|A(\mathbf{x}) - A(\mathbf{y})| \le \|\mathbf{x} - \mathbf{y}\|_1 = |x_1 - y_1| + |x_2 - y_2|).$$

In [2] we have shown the next result.

Theorem 9. *Let $A : [0,1]^2 \to [0,1]$ be a modular 1–Lipschitz aggregation function. Then the function $\tilde{A} : [0,1]^2 \to [0,1]$ given by*

$$\tilde{A}(x,y) = \min(x, y, A(x,y)) \tag{4}$$

is a copula.

Note that, considering any copula $C : [0,1]^2 \to [0,1]$, the function $A_C : [0,1]^2 \to [0,1]$ given by

$$A_C(x,y) = \frac{C(x,x) + C(y,y)}{2}$$

satisfies the constraints of Theorem 9, and then $\tilde{A}_C$ given by (4) is a diagonal copula introduced in [17]. For some further generalizations of Theorem 9 we recommend our paper [2].

We briefly recall some other recent construction methods for copulas. Conic copulas were introduced in [6]. A close relation between Archimax copulas [1] and conic copulas [6] was recently shown in [3]. Interesting, especially for fitting purposes, seems to be also the approach based on perturbations of particular copulas [14]. Finally, recall DUCS copulas (Distorted Univariate Conditioning Stable Copulas) introduced and discussed in [13].

Acknowledgment. The work on this paper was supported by the grant VEGA 1/0420/15.

References

[1] C. Alsina, M.J. Frank, B. Schweizer, *Associative Functions on Intervals: A Primer of Triangular Norms*, World Scientific, Singapore, 2006.
[2] B. De Baets, H. De Meyer, J. Kalická, R. Mesiar, *On the relationship between modular functions and copulas*, Fuzzy Sets and Systems, in press.
[3] M. Dibala, L. Vavríková, *The relations between conic and Archimax copulas*. Submitted.
[4] A. Dolati, M. Úbeda–Flores, *Constructing copulas by means of pairs of order statistics*, Kybernetika **45** (2009), 992–1002.
[5] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, *Aggregation Functions: Encyclopedia of Mathematics*. Cambridge University Press, 2009.
[6] T. Jwaid, B. De Baets, J. Kalická, R. Mesiar, *Conic aggregation functions*, Fuzzy Sets and Systems **167** (2011), pp. 3–20.
[7] A. Khoudraji, *Contributions 'a l'étude des copules et 'a la mod élisation de valeurs extrémes bivariées*, Doctoral Dissertation Thesis, Université Laval, Québec, 1995.
[8] E.P. Klement, A. Kolesárová, R. Mesiar, S. Saminger–Platz, *On the role of ultramodularity in the construction of binary copulas*. Submitted manuscript.
[9] E.P. Klement, M. Manzi, R. Mesiar, *Ultramodular aggregation functions*, Information Sciences **181 (19)** (2011), pp. 4101–4111.
[10] E.P. Klement, M. Manzi, R. Mesiar, *Ultramodularity and copulas*, Rocky Mountain J. Math. **44** (2014), pp. 189–202.
[11] A. Kolesárová, G. Mayor, R. Mesiar, *Quadratic constructions of copulas*, Information Sciences, in press.
[12] E. Liebscher, *Construction of asymmetric multivariate copulas*, J.Multivar. Anal. **99** (2008), pp. 2234–2250.
[13] R. Mesiar, M. Pekárová, Ducs copulas, Kybernetika **46 (6)** (2010), pp. 1069–1077.
[14] R. Mesiar, J. Komorník, M. Komorníková, *Perturbation of bivariate copulas*, Fuzzy Sets and Systems **268** (2015), pp. 127–140.
[15] R. Moynihan, *On τ_T semigroups of probability distribution functions II*, Aequat. Math. **17** (1978), pp. 19–40.
[16] R.B. Nelsen, *An Introduction to Copulas*, (second edition), Springer, New York, 2006.
[17] R.B. Nelsen, G.A. Fredricks, *Diagonal copula*s, Distributions with Given Marginals and Moment Problems (1997), pp. 121-128.

Factor copulas constructed from stochastic processes

Matthias Scherer
(joint work with G. Bernhart, J.-F. Mai, S. Schenk)

Randomizing the parameter(s) of a given parametric family of univariate random variables is a popular technique to enrich the distribution in concern with additional stochastic properties and to create new probability laws. On a multivariate level, another motivation is to introduce dependence to originally independent objects by means of a joint mixture variable affecting multiple random variables in a similar way. A well-known example are extendible Archimedean copulas that can be interpreted as the survival copulas arising from a two step experiment: Firstly, a positive random variable M is simulated. Second, a sequence of exponential random variables with rate parameter M is drawn independently. Other examples are credit-risk models where a joint (random) default probability $p \in (0,1)$ is used as mixture variable in a sequence of Bernoulli(p) experiments, or loss models for insurance claims based on Poisson-distributed count variables with joint (random) intensities.

Factor models created in this way are popular due to, among others, the following facts: They enjoy a great level of interpretability, they are straightforward to simulate in large dimensions, the dimension of the considered problem is flexible, convenient limit results for large-dimensional random vectors $(X_1, \ldots, X_d)$ for $(d \nearrow \infty)$ are often computable, parametric families of mixture variables imply parametric families of copulas, and extensions beyond conditional iid (i.e. homogeneous one-factor models) are typically easy to find by, e.g., using multiple factors or inhomogeneous marginal laws. Moreover, hierarchical constructions are immediate in many cases, see [10].

Most factor models currently considered – in theory as well as in practice – are based on the aforementioned idea of using random parameters. This ansatz, however, can be extended to more involved random objects, e.g. stochastic processes. Providing more mathematical structure, a famous result by Bruno de Finetti (see [2]) shows that an infinite sequence of random variables $\{X_k\}_{k\in\mathbb{N}}$ on $(\Omega, \mathscr{F}, \mathbb{P})$ is exchangeable if and only if it is conditionally iid, i.e. there exists a sub-σ-algebra $\mathscr{G} \subset \mathscr{F}$ s.t. for all $d \geq 2$:

$$\mathbb{P}(X_1 \leq x_1, \ldots, X_d \leq x_d \,|\, \mathscr{G}) = \prod_{k=1}^{d} \mathbb{P}(X_1 \leq x_k \,|\, \mathscr{G}), \quad x_1, \ldots, x_d \in \mathbb{R}.$$

This is equivalent to the existence of a random distribution function $t \mapsto F_t$ such that conditioned on $\mathscr{G} := \sigma(\{F_t\}_{t\in\mathbb{R}})$, the components $\{X_k\}_{k\in\mathbb{N}}$ are iid and can be represented as

$$X_k := \inf\{t \in \mathbb{R} : F_t \geq U_k\}, \quad \{U_k\}_{k\in\mathbb{N}} \overset{\text{iid}}{\sim} \mathscr{U}[0,1],$$

where the sequence $\{U_k\}_{k\in\mathbb{N}}$ is independent of $\{F_t\}_{t\in\mathbb{R}}$.

In the case of randomized parameters, the stochastic process $\{F_t\}_{t\in\mathbb{R}}$ is an ordinary distribution function with random parameters. Opposed to this quite

simplistic approach, we will need "true" stochastic processes in the following. Motivated especially by applications in portfolio credit-risk modeling and insurance, our research group has considered the questions:

- Given an interesting class of multivariate probability distributions, can we identify the conditionally iid subfamily, i.e. the subfamily representable by a one-factor construction?
- How can we create new probability laws starting from one-factor models?

These questions are flanked by the search for stochastic representations, efficient sampling strategies, extensions to multi-factor models, and real-world applications. It is quite difficult to give a general description on how the above questions can be solved, since the answer heavily depends on the family one has in mind. Generally speaking, one first has to identify the exchangeable subfamily of the distribution in concern in dimension d (since this is a necessary condition for extendibility), then one has to let the dimension go to infinity, and finally one has to guess a stochastic model for $\{X_k\}_{k\in\mathbb{N}}$ that ultimately reveals the nature of $\mathscr{G} := \sigma(\{F_t\}_{t\in\mathbb{R}})$.

Since the applications we have in mind involve the modeling of default times or the arrival times of insurance claims, it is not surprising that we consider fatal shock models on the one hand, and various multivariate extensions of the exponential law on the other hand. Fatal shock models have been thoroughly studied and applied in different fields, e.g., finance, hydrology, insurance, and reliability theory. In their classical stochastic representation, distinct shocks E_I hit combinations of components of a d-dimensional vector, i.e. the vector $(X_1, \ldots, X_d)$ is defined as

$$X_k := \min_{\emptyset\neq I\subset\{1,\ldots,d\}} \{E_I : k \in I\}, \quad k = 1, \ldots, d. \tag{1}$$

The seminal example for both – multivariate exponential laws and fatal shock constructions as in (1) – is the Marshall–Olkin law (see [12]), in which the shocks E_I are independent and exponentially distributed. While this is convenient to interpret and use in low dimensions, the number of involved shocks increases exponentially in d, preventing such models from being applicable even in moderate dimensions. This problem is overcome as soon as a one-factor subfamily is identified. For the Marshall–Olkin law, one can show that the exchangeable subfamily is parameterized by d-monotone vectors, which for $(d \nearrow \infty)$ provides a link to completely monotone sequences. These are linked to Bernstein functions by a result in [4], which are finally in a one-to-one relation with Lévy subordinators (see [9]). The model for $\{X_k\}_{k\in\mathbb{N}}$, called Lévy-frailty construction in [8], can then be formulated as

$$X_k := \inf\{t \geq 0 : \Lambda_t \geq \epsilon_k\}, \quad k \in \mathbb{N}, \tag{2}$$

with a sequence of iid unit exponentials $\{\epsilon_k\}_{k\in\mathbb{N}}$ and $\{\Lambda_t\}_{t\geq 0}$ a Lévy subordinator. On a theoretical basis, this provides a marvelous link between classes of multivariate probability laws and the corresponding families of stochastic processes.

Generalizations to arbitrary exchangeable fatal shock models, extending the Marshall–Olkin model beyond the embedded exponential law, are considered in

[5]. Among others, all copulas of the functional form

$$C(u_1, \ldots, u_d) = \prod_{k=1}^{d} g_k(u_{(k)}), \quad d \geq 2, \tag{3}$$

are characterized analytically and by means of a fatal shock model. Here, $u_{(1)} \leq \ldots \leq u_{(d)}$ denotes the ordered arguments of C. Moreover, a link to additive processes, serving as stochastic factor, is revealed, and the specific cases of Sato processes (see [6]) and the Dirichlet process (see [7]) are discussed in quite some detail.

Turning to alternative multivariate definitions of the exponential law (apart from the Marshall–Olkin distribution), MSMVE distributions and distributions with exponential minima are natural candidates, see [3]. For both classes, the extendible subfamilies are characterized in [11] and a stochastic model based on strong (respectively weak) IDT subordinators is given. Specific families of MSMVE distributions (respectively their associated extreme-value copula) are constructed in [1]. This provides new (low-parametric) families of extreme-value copulas that might be interesting for applications, among others, due to their convenient stochastic representation.

References

[1] Bernhart, G.; Mai, J.-F.; Scherer, M.: *On the construction of low-parametric families of min-stable multivariate exponential distributions in large dimensions.* Dependence Modeling, **3** (2015) 29–46.

[2] De Finetti, B.: *La prévision: ses lois logiques, ses sources subjectives.* Annales de l'Insitut Henri Poincaré, **7** (1937) 1–68.

[3] Esary, J.D.; Marshall, A.W.: *Multivariate distributions with exponential minimums.* Annals of Statistics, **2**(1), (1974) 84–98.

[4] Gnedin, A.; Pitman, A.: *Moments of convex distribution functions and completely alternating sequences.* IMS Collections, Probability and Statistics: Essays in Honor of David A. Freedman, **2**, (2008) 30–41.

[5] Mai, J.-F.; Schenk, S.; Scherer, M.: *Exchangeable exogenous shock models.* Forthcoming in Bernoulli, (2015).

[6] Mai, J.-F.; Schenk, S.; Scherer, M.: *Two novel characterizations of self-decomposability on the half-line.* Working paper, (2015).

[7] Mai, J.-F.; Schenk, S.; Scherer, M.: *Analyzing model uncertainty via a distortion of the stochastic root: a Dirichlet prior approach.* Working paper, (2015).

[8] Mai, J.-F.; Scherer, M.: *Lévy-frailty copulas*, Journal of Multivariate Analysis **100**(7), (2009) 1567–1585.

[9] Mai, J.-F.; Scherer, M.: *Reparameterizing Marshall–Olkin copulas with applications to sampling.* Journal of Statistical Computation and Simulation, **81**(1) (2011) 59–78.

[10] Mai, J.-F.; Scherer, M.: *H-extendible copulas*, Journal of Multivariate Analysis **110**, (2012) 151–160.

[11] Mai, J.-F.; Scherer, M.: *Characterization of extendible distributions with exponential minima via processes that are infinitely divisible with respect to time*, Extremes, **17**(1) (2014), 77–95.

[12] Marshall, A.W.; Olkin, I.: *A multivariate exponential distribution.* Journal of the American Statistical Association, **62** (1967) 30–44.

Reporter: Alexander Ristig

Participants

Prof. Dr. Elif Acar
Department of Statistics
University of Manitoba
338 Machray Hall
Winnipeg, MB R3T 2N2
CANADA

Dr. Valeria Bignozzi
Department of Methods and Models for Economics, Territory and Finance
University of Rome Sapienza
Via del Castro Laurenziano 9
00161 Roma
ITALY

Dr. Axel Bücher
Fakultät für Mathematik
Ruhr-Universität Bochum
Universitätsstr. 150
44801 Bochum
GERMANY

Prof. Dr. Cathy Yi-Hsuan Chen
Department of Finance
Chung-Hua University
No. 707, Section 2, Xiangshan District
WuFu Rd.
Hsinchu City 300
TAIWAN

Prof. Dr. Claudia Czado
Lehrstuhl f. Mathematische Statistik
Technische Universität München
85747 Garching bei München
GERMANY

Prof. Dr. Holger Dette
Fakultät für Mathematik
Ruhr-Universität Bochum
44780 Bochum
GERMANY

Prof. Dr. Fabrizio Durante
Faculty of Economics & Management
Free University of Bozen-Bolzano
Piazza Università, 1
39100 Bolzano
ITALY

Prof. Dr. Jean-David Fermanian
CREST - ENSAE
Centre Recherche de l'Economie et de la Statistique
15, Boulevard Gabriel
92245 Malakoff Cedex
FRANCE

Prof. Dr. Anne-Laure Fougères
Institut Camille Jordan
Université Claude Bernard Lyon 1
43 blvd. du 11 novembre 1918
69622 Villeurbanne Cedex
FRANCE

Prof. Dr. Christian Genest
Department of Mathematics & Statistics
McGill University
805, Sherbrooke Street West
Montreal, P.Q. H3A 2K6
CANADA

Prof. Dr. Irène Gijbels
Department of Mathematics and
Center for Statistics
Katholieke Universiteit Leuven
W. de Croylaan 54
3001 Leuven (Heverlee)
BELGIUM

Prof. Dr. Wolfgang K. Härdle
Wirtschaftswissenschaftliche Fakultät
Ladislaus v. Bortkiewicz Chair of Statistics CASE
Humboldt-Universität zu Berlin
Unter den Linden 6
10117 Berlin
GERMANY

Prof. Dr. Marius Hofert
Department of Statistics & Actuarial Science
Faculty of Mathematics
University of Waterloo
Waterloo, Ontario N2L 3G1
CANADA

Prof. Dr. Piotr Jaworski
Institute of Mathematics
University of Warsaw
ul. Banacha 2
02-097 Warszawa
POLAND

Prof. Dr. Alexander J. McNeil
Dept. of Actuarial Mathematics & Statistics
Heriot-Watt University
Edinburgh EH14 4AS
UNITED KINGDOM

Prof. Dr. Radko Mesiar
Department of Mathematics
Faculty of Civil Engineering
Slovak University of Technology
Radlinskeho 11
81005 Bratislava
SLOVAKIA

Prof. Dr. Cathy Ning
Department of Economics
Ryerson University
350 Victoria Street
Toronto Ont. M5B 2K3
CANADA

Prof. Dr. Ostap Okhrin
Wirtschaftswissenschaftliche Fakultät
Lehrstuhl für Statistik
Humboldt-Universität zu Berlin
Spandauer Strasse 1
10178 Berlin
GERMANY

Prof. Dr. Bruno Remillard
HEC Montreal
3000, chemin de la
Cote-Sainte-Catherine
Montreal (Quebec) H3T 2A7
CANADA

Alexander Ristig
Wirtschaftswissenschaftliche Fakultät
Humboldt-Universität Berlin
Spandauer Straße 1
10178 Berlin
GERMANY

Prof. Dr. Louis-Paul Rivest
Dept. de Mathématiques et de Statistique
Université Laval
Cite Universitaire
Quebec, QC G1K 7P4
CANADA

Prof. Dr. Matthias Scherer
Zentrum Mathematik
TU München
Boltzmannstr. 3
85748 Garching b. München
GERMANY

Prof. Dr. Johan Segers
Institut de Statistique
Université Catholique de Louvain
Voie du Roman Pays, 20
1348 Louvain-la-Neuve
BELGIUM

Prof. Dr. Carlo Sempi
Dip. di Matematica e Fisica "E. de Giorgi"
Universita del Salento
C.P. 193
73100 Lecce
ITALY

Prof. Dr. Peter Song
Department of Biostatistics
School of Public Health
University of Michigan
1415 Washington Heights
Ann Arbor, MI 48109-2029
UNITED STATES

Prof. Dr. Johanna F. Ziegel
Institut für Mathematische Statistik
und Versicherungslehre
Universität Bern
Sidlerstrasse 5
3012 Bern
SWITZERLAND